

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:48:20 ; Search time 19.89 Seconds
(without alignments)
391.314 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 498

Sequence: 1 CS0NEYFDSLHACIPCOLR.....SEYFDSLHACPAFCQPYC 81

Scoring table: BIOSYM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	40.4	184	2	B-cell maturation
2	94.5	19.0	223	2	ultra-high-sulfur
3	94.5	19.0	230	2	ultra-high-sulfur
4	92.5	18.6	169	1	ultra-high-sulfur
5	91.5	18.4	186	2	ultra-high-sulfur
6	89	17.9	1574	2	MEGF6 protein - ra
7	87	17.5	1680	2	furin (EC 3.4.21.7
8	86.5	17.4	1548	2	serine proteinase
9	86	17.3	188	2	high sulfur protei
10	83	16.7	131	1	keratin high-sulfu
11	83	16.7	175	2	keratin high-sulfu
12	82	16.5	126	2	keratin high-sulfu
13	82	16.5	526	2	keratin high-sulfu
14	81	16.3	2823	2	keratin high-sulfu
15	81	16.3	2823	2	keratin high-sulfu
16	81	16.3	3102	2	keratin high-sulfu
17	80.5	16.2	151	2	keratin high-sulfu
18	80.5	16.2	937	2	keratin high-sulfu
19	80.5	16.2	1101	2	keratin high-sulfu
20	80.5	16.2	1798	2	keratin high-sulfu
21	80	16.1	3075	2	keratin high-sulfu
22	79.5	16.0	965	2	keratin high-sulfu
23	79.5	16.0	1077	2	keratin high-sulfu
24	79.5	16.0	3133	2	keratin high-sulfu
25	79	15.9	861	2	keratin high-sulfu
26	78.5	15.8	201	2	keratin high-sulfu
27	78.5	15.8	294	2	keratin high-sulfu
28	78	15.7	572	2	keratin high-sulfu
29	78	15.7	1188	2	keratin high-sulfu

30	77.5	15.6	132	1	KRCN3T	keratin high-sulfu
31	77.5	15.6	177	2	S37650	high-sulfur kerati
32	77	15.5	162	2	I47107	high-sulfur wool m
33	77	15.5	172	1	KRSHHA	keratin high-sulfu
34	76.5	15.4	5376	2	T42215	keratin high-sulfu
35	76.5	15.4	1513	2	T52527	keratin high-sulfu
36	76.5	15.4	1513	2	T23681	keratin high-sulfu
37	76.5	15.4	2414	2	A54277	keratin high-sulfu
38	76	15.3	172	2	I47106	keratin high-sulfu
39	76	15.3	182	2	I47105	keratin high-sulfu
40	76	15.3	1299	2	T43251	keratin high-sulfu
41	76	15.3	2440	2	S39162	keratin high-sulfu
42	76	15.3	2441	2	S39161	keratin high-sulfu
43	76	15.3	2910	2	T42214	keratin high-sulfu
44	75.5	15.2	654	2	T30136	keratin high-sulfu
45	75.5	15.2	1713	2	A53347	keratin high-sulfu

ALIGNMENTS

RESULT 1
S43486
B-cell maturation factor - human
N:Alternate names: BCM protein; BCMA protein; BEL protein
C:Species: Homo sapiens (man)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2000
C:Accession: S43486; S31208; S36661
R:Laab, Y.; Gras, M.P.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapts, A.
Nucleic Acids Res. 22, 1147-1154, 1994
A>Title: The BCM gene, preferentially expressed during B lymphoid maturation, is bid
A:Reference number: S43486; M0ID:94218235
A:Accession: S43486
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-184 <LAA>
A:Cross-references: EMBL:229574; NID:9471244; PIDN:CAA82690.1; PID:9471245
R:Laab, Y.; Gras, M.P.; Carbonnel, F.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapi
EMBO J. 11, 3897-3904, 1992
A>Title: A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t
A:Reference number: S31208; M0ID:93010984
A:Accession: S31208
A:Molecule type: mRNA
A:Residues: 1-184 <LAA>
A:Cross-references: EMBL:214954; NID:929407; PIDN:CAA78679.1; PID:929408
A:Accession: S36661
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 4-184 <LAA>
A:Cross-references: EMBL:214955
C:Genetics:
A:Gene: GDB:BCMA
A:Cross-references: GDB:135977; OMIM:109545
A:Map position: 16p13.1-16p13.1
A:Introns: 44/1; 93/71
C:Superfamily: human B-cell maturation factor

Query Match 40.4%; Score 201; DB 2; Length 184;
Best Local Similarity 100.0%; Pred. No. 8.1e-11;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CS0NEYFDSLHACIPCOLRCSNTPPLTCQRYC 34
DB 8 CS0NEYFDSLHACIPCOLRCSNTPPLTCQRYC 41

RESULT 2
B38346
ultra-high-sulfur keratin 2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 24-Sep-1999
C:Accession: A38660; B38346

Db 85 CQPCGQSCCQSSCCGPRCCSSCCGPRCCISSCCGPC--CRSCCQSSC----- 132
QY 70 HACPPATCQPYC 81
| | ||:
Db 133 --CRPC-CQPC 141 *

RESULT	6
T13954	

C:Species:	Rattus norvegicus (Norway rat)
C:Date:	20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession:	T133954
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.	
R:Genomics	51, 27-34, 1998
A:Title:	Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number:	Z4126; MUID:98360089
A:Accession:	T133954
A:Status:	preliminary; translated from GB/EMBL/DBJ
A:Molecule type:	mRNA
A:Residues:	1-1574 <NNA>
A:Cross-references:	EMBL:AB011532; NID:g3449293; PIRDB:BA03462.1; PID:g3449294
A:Experimental source:	strain Sprague-Dawley; brain
C:Genetics:	
C:Gene:	MEGF6

Query Match	17.9%	Score 89;	DB 2;	Length 1574;
Best Local Similarity	29.4%;	Pred. No. 1.8;		
Matches 30;	Conservative 9;	Mismatches 25;	Indels 38;	Gaps 8;

```

QY      1 CSONEFYFSLHACIPCOL-----RCSSNPPLT-----CQRCCCEY----FSLHACPC 47
      || :: :|||  || : |||  | : | : |||  |
Db      966 CSAGAPCDAVTSCI-CPAGRMGPACQAQSCPPLTFGLNCSQICTCFNGASCDSVTGQC-- 1022

```

```

OY      48 LRCSP-----PTCQYCCFHSEYFDLSLHACPPA-----TCQPYC 81
          | : | | | | | | | | | | | | | | | | | | | |
DB      1023 -HCAPGWMGPTC-----LQACRPGLYGKNCQHS 1050

```

RESULT	7
A43434	

I:turn (BC 3.4.21.75) 2 - fruit fly (*Drosophila melanogaster*)
 C:Species: *Drosophila melanogaster*
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 31-Mar-2000
 C:Accession: A43434
 R:Rehbroek, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dumke, U.; Rentrop, M.; Gateff, E.
 J. Biol. Chem. 267, 17208-17215, 1992
 A:Title: Cloning and functional expression of Dfurn12, a subtilisin-like proprotein proc
 A:Reference number: A43434; MUID:92381036
 A:Accession: A43434
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1680 <R0>
 A:Cross-references: GB|M94375; NID:g157461; PID:g157462.
 A:Note: sequence extracted from NCBI backdome (NCBIN:111933, NCBIP:111934)
 C:Genetics:
 A:Gene: FlyBase:Fur2
 A:Cross-references: FlyBase:Fpqn0004598
 C:Superfamily: subtilisin homology
 C:Keywords: hydrolase; serine protease; transmembrane protein
 E:409-652/domain: subtilisin homology <SBT>
 E:418,457,638/Active site: Asp, His, Ser #status predicted

Query Match	17.5%	Score	87	DB	2	Length	1680
Best Local Similarity	25.0%	Pred. No.	2.9				
Matches	31	Conservative	12	Mismatches	31	Indels	50
						Gaps	9

```

Oy      1 CSONEYFDSLLHACIPQOLRCSSNTPPLTQ--RYC-CEYFDSLH-----ACP--- 46
      | : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 1051 CPDGYFEENRNRLCYVCEPCNCA-----QDHPREYCTSCDH-HLVNHEHKCSACP.LDT 1103

```

```

Oy  47  -----CLRGSPPTCOYC--CPHSYE--DSLILHACPA-----TC 77
      | | | | | | | | | | | | | | | | | | | | | | | |
Db  1104 YETEDNKCACFCHSTCATCNGPTDDCLTCSSRYANQNKCLISCPGFIADKKRLECAPC 1163
Oy      78  QPYC  81
      | |
Db  1164 QEGC  1167

```

QY	78	QPYC	81
Db	1164	QEGC	1167

RESULT 8
S34583

C:Accession: S34583
C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
R:Nakagawa, T.; Murakami, K.; Nakayama, K.
FEBS Lett. 327, 165-171, 1993
A:Title: Identification of an isoform with an extremely large Cys-rich region of PC6,
A:Reference number: S34583; MUID:93327934
A:Accession: S34583
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1548 <NA>
A:Cross-references: GB:D17583; NID:g407344; PIDN:BA04507.1; PID:d1005033; PID:g440337
C:Keywords: hydrolase; serine proteinase

Query Match	17.4%	Score	86.5	DB	2	length	1548
Best Local Similarity	22.1%	Pred	No. 3				
Matches	29	Conservative	6	Mismatches	43	Indels	53
						Gaps	3

```

Qy      1  CSQNEYFDLSLHACIPQQLKRCSSNTPPLTCQRYCCETFDLSL----- 42
      1  :  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      1152 CAAYEYWDGSHRCQPCHKKCSRCSPSEDCYCTCPRETFLLNTTCVKCEPEGYHTDKS 1211

```

```

Qy 43 -----HACPCLKSP-----PTQYCCFHSEYFDLSLHACP-----73
      | : | | | | | | | |
Db 1212 QQCVLCHSSCRPCGPHSMQCLSCRPGWFYLGKECLLDQRDGYGGESTSGCEKCDKSK 122

```

```
QY      74  -----PATCQ  78
          |  |
Db  1272  SCRGPRPTDCQ 1282
```

Db 1272 SCRGPRPTDQ 1282

RESULT 9
JC6547

high sulfur protein B2E - rat
C.Species: Rattus norvegicus (Norway rat)
C.Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jul-2000
C.Accession: J06547
R.Mitsui, S.; Ohuchi, A.; Adachi-Yamada, T.; Hotta, M.; Tsuboi, R.; Ogawa, H.
Gene 208, 123-129, 1998
A.Title: Structure and hair follicle-specific expression of genes encoding the rat h
A.Reference number: J06547; MUID:98201605
A.Accession: J06547
A.Molecule type: DNA
A.Residues: 1-188 <MTR>
A.Cross-references: DDBJ:AB003753; NID:93046870; PIDN:BAA5573.1; PID:93046871
C.Comment: This protein is a cysteine-rich, keratin associated protein.
C.Genetics:
A.Gene: b2E
C.Superfamily: keratin high-sulfur matrix protein IITA
C.Keywords: hair

Query Match	17.3%	Score 86;	DB 2;	Length 188;
Best Local Similarity	22.58;	Pred. No. 0.73;		
Matches 29; Conservative	10;	Mismatches 28;	Indels 62;	Gaps 7;

```

QY      1 CSQNEYFDLSLHACIP-----CQLRCSSNTPPLTCQ-----R 32
      |||:  | | ||| | | : |||
Db 69 CSQS-----CCQPSCCQTSCCQPTCCQNS-----SCQTSCCGTSGGQEGSGATSCVR 118

```


Search completed: May 15, 2002, 09:49:35
Job time: 75 sec

A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2823 <W1L>
A:Cross-references: EMBL:AL008585; PIDN:CAAI5432.1; GSPDB:GN00019; CESP:T22A3.8
A:Experimental source: clone W10E24
R:McMurray, A.
submitted to the EMBL
A:Reference number: Z19980
A:Accession: T25096
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2823 <W12>
A:Cross-references: EMBL:281125; PIDN:CAB03385.1; GSPDB:GN00019; CESP:T22A3.8
A:Experimental source: clone T22A3
C:Genetics:
A:Gene: CESP:T22A3.8
A:Map position: 1
A:introns: 45/1; 282/2; 312/3; 416/2; 1255/3; 1329/3; 1418/3; 1776/2; 1988/2; 2760/2
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match
Best Local Similarity 27.7%; Score 81; DB 2; Length 2823;
Matches 26; Conservative 10; Mismatches 32; Indels 26; Gaps 7;

QY 1 CSQNEFYDSLHACIPCOLRCSSNTPPL--TCQRYCCEYFDSLHACPCLCRC----SPP 53
DB 873 CSQGFEDPLTGKCIKC--TCNGNIDPMGIGNC-----DS--ETGKCLKICGHTTGD 920
QY 54 TCQYCCFHSEYFDSLHACPRATC-----QPYC 81
DB 921 SCSECKEH-HMGNAQLHTCKPCGCHTQGA VNPQC 953

RESULT 15

F87908
Protein T22A3.8 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 24-Aug-2001
C:Accession: F87908; E87908
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998

A:title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A78000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Accession: F87908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2823 <STO>
A:Cross-references: GB:chr_1; PIDN:CAAI5432.1; PID:g3924779; GSPDB:GN00019; CESP:T22A3.8
A:Accession: E87908
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2823 <ST>
A:Cross-references: GB:chr_1; PIDN:CAB03385.1; PID:g3924881; GSPDB:GN00019; CESP:T22A3.8
C:Genetics:
A:Gene: T22A3.8
A:Map position: 1
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match
Best Local Similarity 27.7%; Score 81; DB 2; Length 2823;
Matches 26; Conservative 10; Mismatches 32; Indels 26; Gaps 7;

QY 1 CSQNEFYDSLHACIPCOLRCSSNTPPL--TCQRYCCEYFDSLHACPCLCRC----SPP 53
DB 873 CSQGFEDPLTGKCIKC--TCNGNIDPMGIGNC-----DS--ETGKCLKICGHTTGD 920
QY 54 TCQYCCFHSEYFDSLHACPRATC-----QPYC 81
DB 921 SCSECKEH-HMGNAQLHTCKPCGCHTQGA VNPQC 953

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:49:40 ; Search time 10.13 Seconds
(without alignments)
309,604 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 496
Sequence: 1 CSQNEFDSLILHACIPQQLR.....SEYFDSLILHACPATCPYPC 81

Scoring table: BLOSUM62 ;
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	40.4	184	TR17_HUMAN	002223 homo sapien
2	92.5	18.6	163	KRUB_HUMAN	P26371 homo sapien
3	91	18.3	194	KRUB_HUMAN	075690 homo sapien
4	87	17.5	1680	FUR2_DROME	P30432 drosophila
5	86.5	17.4	1877	PKC5_MOUSE	004592 homo muscu
6	83	16.7	131	KRA3_SHEEP	P02441 ovine arles
7	82	16.5	526	UBPW_MOUSE	061068 homo muscu
8	81	16.3	1587	LMG3_HUMAN	094606 homo sapien
9	80.5	16.2	937	PAC4_RAT	063415 ratulus norv
10	80.5	16.2	1798	LMB2_HUMAN	P55268 homo sapien
11	80	16.1	3075	LMG1_HUMAN	P25391 homo sapien
12	79.5	16.0	965	YNC3_YEAST	P53921 saccharomyc
13	79.5	16.0	3133	HMCT_BOMO	P98092 bombyx mori
14	79	15.9	1696	PKC5_BRACL	09n115 brachioosto
15	77.5	15.6	132	KRA3_CAVHI	P02442 capra hircu
16	77.5	15.6	139	KRA3_CAVHI	043609 capra hircu
17	77	15.5	171	KRA3_SHEEP	P02438 ovine arles
18	77	15.5	5376	ZAN_MOUSE	088799 mus muscu
19	76.5	15.4	2414	CBP_MOUSE	P45481 mus muscu
20	76.5	15.4	2441	CBP_MOUSE	016787 mus muscu
21	75.5	15.2	1713	LMG3_HUMAN	0910b6 mus muscu
22	75	15.1	1581	LMG3_MOUSE	P11047 homo sapien
23	74.5	15.0	1609	LMG1_HUMAN	028065 bos tauris
24	74.5	15.0	610	CABP_BOVIN	028065 bos tauris
25	74	14.9	1786	LMB1_HUMAN	P07942 homo sapien
26	74	14.9	2442	CBP_HUMAN	092793 homo sapien
27	73.5	14.8	313	SPY1_MOUSE	0949v9 mus muscu
28	73.5	14.8	1799	LMB2_MOUSE	061992 mus muscu
29	72.5	14.6	969	PAC4_HUMAN	P29132 homo sapien
30	72.5	14.6	1592	SORL_CHICK	098930 g scortilin-
31	72	14.5	575	TRBM_HUMAN	P07204 homo sapien
32	72	14.5	867	SPRO_BOVIN	P98167 bos tauris
33	72	14.5	1367	IGIR_HUMAN	P08069 homo sapien

34	72	14.5	1700	BAR3_CHITE	003376 chironomus
35	72	14.5	2569	LMG3_MOUSE	061789 mus muscu
36	72	14.5	4655	LRP2_HUMAN	P98164 homo sapien
37	71.5	14.4	1435	EBAL_PLARC	P19214 plasmodium
38	71	14.3	62	IBB_MEDSC	P80321 medicago sc
39	71	14.3	63	ICEL_ASCSU	P07851 ascaris suu
40	71	14.3	810	NEIL_HUMAN	092832 homo sapien
41	71	14.3	3712	LMG1_DROME	000174 drosophila
42	70.5	14.2	96	IBBA_PEA	041065 pisum sativ
43	70.5	14.2	1104	NEFX_HUMAN	012986 homo sapien
44	70	14.1	181	KR2D_SHEEP	P08131 ovine arles
45	70	14.1	3110	LMG2_HUMAN	P24043 homo sapien

ALIGNMENTS

RESULT	ID	TR17_HUMAN	STANDARD	PRT	184 AA.
AC	002223				
DT	01-JUL-1993	(Rel. 26, Last Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 17 (B-cell maturation protein).				
GN	TNFRSF17 OR BCM4 OR BCM.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.				
RC	TISSUE-Peripheral blood leukocytes, and lymph node;				
RX	MEDLINE=93010984; PubMed=1396583;				
RA	Laabl Y., Gras M.P., Cardonnel F., Brouet J.C.; Berger R.,				
RA	Larsen C.J., Tsapis A.;				
RT	"A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t(4;16)(q26;p13) translocation in a malignant T cell lymphoma.",				
RL	EMBO J. 11:3897-3904(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94218235; PubMed=8165126;				
RA	Laabl Y., Gras M., Brouet J., Berger R., Larsen C., Tsapis A.;				
RT	"The BCM4 gene, preferentially expressed during B lymphoid maturation, is bidirectionally transcribed.";				
RL	Nucleic Acids Res. 22:1147-1154(1994).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99425270; PubMed=10493829;				
RA	Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,				
RA	Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,				
RA	Destalles Mays A., Gao Y., Xu R.X., Kang H.-L., Mitchell S.,				
RT	Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;				
RL	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";				
RN	Genomics 60:295-308(1999).				
RN	[4]				
RP	SEQUENCE FROM N.A., AND VARIANT THR-153.				
RA	Kawasaki A., Tsuchiya N., Fukazawa T., Hashimoto H., Tokunaga K.;				
RT	"New polymorphisms of human BCM4.";				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	FUNCTION.				
RX	MEDLINE=20363816; PubMed=10903733;				
RA	Hatzoglou A., Roussel J., Bourgeade M.F., Rogier E., Madry C.,				
RA	Inoue J., Devergne O., Tsapis A.;				
RT	"TNF receptor family member BCM4 (B cell maturation) associates with TNF receptor-associated factor (TRAF) 1, TRAF2, and TRAF3 and activates NF-kappa B, elk-1, c-Jun N-terminal kinase, and p38				
RT	mitogen-activated protein kinase.";				
RL	J. Immunol. 165:1322-1330(2000).				
RN	[6]				

```

RP FUNCTION.
RA MEDLINE-20259066; PubMed-10801128;
RA Gross J.A., Johnston J., Mudri S., Enselman R., Dillon S.R.,
RA Madden K., Xu W., Parrish-Novak J., Foster D., Lofton-Day C.,
RA Moore M., Litlau A., Grossman A., Haugen H., Foley K., Blumberg H.,
RA Harrison K., Kindsvogel W., Clegg C.H.;
RA "TACI and BCMA are receptors for a TNF homologue implicated in B-cell
RA autoimmune disease.";
RT Nature 404:995-999(2000).
RL
CC -1- FUNCTION: Receptor for TNFSF13B/BLYS/BAFF.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE B-CELLS.
CC -1- DISEASE: A FORM OF T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA (T-ALL) IS
CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(4;16)(Q26;P13)
CC WHICH INVOLVES BCMA AND IL2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z14954; CAA78679.1; -
DR EMBL: Z29575; CAA82691.1; -
DR EMBL: Z29574; CAA82690.1; -
DR EMBL: U95742; AAB67251.1; -
DR EMBL: AB052772; BAB60895.1; -
DR PIR: S31208; S31208.
DR PIR: S31209; S31209.
DR MIM: 109345; -
RW Receptor; Proto-oncogene; Transmembrane; Chromosomal translocation;
KW Polymorphism.
FT SITE 3 4 POTENTIAL.
FT TRANSMEM 55 77 BREAKPOINT FOR TRANSLOCATION TO FORM
FT INTERLEUKIN 2/BCM ONCOGENE.
FT VARIANT 153 153 A->T.
FT SEQUENCE 184 AA; 20138 MW; 277AF1E2767D932 CRC64;
SQ
Query Match 40.4%; Score 201; DB 1; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.9e-13;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CSONEYFDSLHACIPCOLRCSSNTPPLTCORYC 34
DB 8 CSONEYFDSLHACIPCOLRCSSNTPPLTCORYC 41

```

```

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99148005; PubMed-10023043;
RA Perez C., Auriol J., Gerst C., Bernard B.A., Egly J.-M.;
RT "Genomic organization and promoter characterization of two human UHS
RT keratin genes.";
RL Gene 227:137-148(1999).
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -1- TISSUE SPECIFICITY: CUTICLE LAYERS OF DIFFERENTIATING
CC HAIR FOLLICLES.
CC -1- DEVELOPMENTAL STAGE: AT A LATE STAGE OF FIBER DIFFERENTIATION.
CC -1- DOMAIN: MAINLY COMPOSED OF CYS-RICH (CR), GLY-RICH (GR) AND SER-
CC RICH (SR) REPEATS.
CC -1- SIMILARITY: BELONGS TO THE UHS KERATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X55293; CAA39005.1; -
DR EMBL: AJ006693; CAA07189.1; -
DR HSSP: P04355; 2MRT.
DR MIM: 148021; -
KW Keratin; Repeat; Multigene family.
FT SEQUENCE 169 AA; 16276 MW; 219B14FEEB49D4A.CRC64;
SQ
Query Match 18.6%; Score 92.5; DB 1; Length 169;
Best Local Similarity 30.0%; Pred. No. 0.012;
Matches 27; Conservative 5; Mismatches 37; Indels 21; Gaps 5;
QY 1 CSONEYFDSLHACIPCOLRCSSNTPPLTCORCYCEYFDSLHAC-PCLR-----CS 51
DB 77 CSQ-----CSCCKPFC--CCSSGCGSSCCOCCKPFCSCQSCCKPCCSSGGRSSCCQ 127
QY 52 PPTCYCCFHFSEYFDSLHACPPATCPQYC 81
DB 128 SSCCKPCCSSGCGSS---CCQSSCKPCC 154

```


CC 6-20 kDa, WHEREAS THE MICROFIBRILLS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC
CC -I- TISSUE SPECIFICITY: CUTICLE LAYERS OF DIFFERENTIATING
CC HAIR FOLLICLES.
CC
CC -I- DOMAIN: MAINLY COMPOSED OF CYS-RICH (CR), GLY-RICH (GR) AND SER-
CC RICH (SR) REPEATS.
CC
CC -I- SIMILARITY: BELONGS TO THE UNS KERATIN FAMILY.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AJ006692; CAA07188.1; -
DR HSSD: P04355; 2MRP.
KW Keratin; Repeat; Multigene family.
SQ SEQUENCE 194 AA: 18094 MW: 8607B2AE906ED4A CRC64;

RESULT 5
PC5_MOUSE STANDARD: PRT: 1877 AA.
AC 004592: 062040;
DT 01-FEB-1995 (rel. 31, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)
DE (Proteinase PC5) (PC5) (Subtilisin/kexin-like protease PC5)
DE (Proteinase PC5) (PC5) (Subtilisin-like proprotein convertase 6)
DE (SPC6).
GN PCSK5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RN SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PC5B).
RC STRAIN=ICR; TISSUE=Intestine;
RX MEDLINE=93327934; PubMed=8335106;
RA Nakagawa T., Murakami K., Nakayama K.;
RT "Identification of an isoform with an extremely large Cys-rich region
of PC6, a kex2-like processing endoprotease.";
RT FEBS Lett. 327:165-171(1993).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Brain, and Intestine;
RX MEDLINE=93224489; PubMed=8468318;
RA Nakagawa T., Hosaka M., Torii S., Watanabe T., Murakami K.,
RA Nakayama K.;
RT "Identification and functional expression of a new member of the
mammalian kex2-like processing endoprotease family: its striking
structural similarity to PACE4.";
RT J. Biochem. 113:132-135(1993).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORM PC5A).
RC TISSUE=Adrenal cortex;
RX MEDLINE=93342056; PubMed=8341687;
RA Lussan J., Vleau D., Hamelin J., Day R., Chretien M., Seidah N.G.;
RT "cDNA structure of the mouse and rat subtilisin/kexin-like PC5: a
candidate proprotein convertase expressed in endocrine and
nonendocrine cells.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).
RN [4]
RN PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.
RX MEDLINE=97103178; PubMed=8947550;
RA De Bie I., Marchkiewicz M., Malide D., Lazure C., Nakayama K.,
RA Bendayan M., Seidah N.G.;
RT "The isoforms of proprotein convertase PC5 are sorted to different
subcellular compartments.";
RT J. Cell Biol. 135:1261-1275(1996).
RN [5]
RN DEVELOPMENTAL EXPRESSION.
RX MEDLINE=96293353; PubMed=8698813;
RA Constant D.B., Calfon M., Robertson E.J.;
RT "SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone
morphogenetic proteins at distinct sites during embryogenesis.";
RL J. Cell Biol. 134:181-191(1996).
RN [6]
RN DEVELOPMENTAL EXPRESSION.
RX MEDLINE=97436919; PubMed=9291583;
RA Rancourt S.L., Rancourt D.E.;
RT "Murine subtilisin-like proteinase SPC6 is expressed during embryonic
implantation, somitogenesis, and skeletal formation.";
RL Dev. Genet. 21:75-81(1997).
RN [7]
RN FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY
WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE
OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE
FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED
IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION
OF GROWTH FACTORS.

CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR
CC PROPROTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS, WHERE XAA
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.
CC -1- SUBCELLULAR LOCATION: PC5A IS SECRETED THROUGH THE REGULATED
CC SECRETORY PATHWAY. PC5B IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH
CC EARLY ENDOSOMES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PC5B/LONG (SHOWN HERE)
CC AND PC5A/SHORT; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: PC5A IS EXPRESSED IN MOST TISSUES BUT IS MOST
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PC5B IS EXPRESSED IN THE
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,
CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE LIVER AND THE LIVER,
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT
CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT
CC E7.5, INTENSE EXPRESSION IN EXTRAEMBRYONIC ENDODERM, AMNION AND
CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK
CC SAC FOLLOWED BY A CONFINATION TO DERMATOTOME COMPARTMENT. BETWEEN
CC E9.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL
CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED
CC TO THE CONDENSING MESenchyma SURROUNDING THE CARTILAGE. AT THIS
CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL
CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E16.5,
CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.
CC ISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF
CC ISOFORM B OCCUR AT E12.5.
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE
CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC
CC RETICULUM.
CC -1- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN
CC SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS
CC WITH THE TGN SORTING PROTEIN PACS-1.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SIMILARITY FAMILY.
CC -1- SIMILARITY: CONTAINS 1 HOWO B/P DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D17583; BAA04507.1; -;
DR EMBL: D12619; BAA02143.1; -;
DR EMBL: L14932; AAA74636.1; -;
DR PIR: JX0248; JX0248.
DR PIR: A48225; A48225.
DR HSSP: Q99405; IMPT.
DR MEROPS: S08.076; -;
DR MGD: MGI:97515; PCSK5.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002174; Furin-like.
DR InterPro: IPR002884; P_domain.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF01483; P; 1.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PRODOM: PD000717; P_domain; 1.
DR SMART: SM00181; EGF_3.
DR SMART: SM00001; EGF-like; 2.
DR SMART: SM00261; FU; 22.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HTS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; signal;
KW Cleavage on pair of basic residues; Repeat; Alternative splicing;
KW Transmembrane.
FT SIGNAL 1 34
FT PROPEP 35 116 PROPROTEIN CONVERTASE SUBTILISIN/KEXIN
FT CHAIN 117 1877

```

FT DOMAIN 117 1768 TYPE 5.
FT TRANSMEM 1769 1789 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1790 1877 POTENTIAL.
FT DOMAIN 117 452 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 464 602 CATALYTIC.
FT DOMAIN 638 1753 HOMO B.
FT DOMAIN 1825 1844 CYS-RICH MOTIF (CRM) REGION.
FT DOMAIN 1856 1877 AC 1.
FT SITE 116 117 AC 2.
FT SITE 521 523 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 173 173 CELL ATTACHMENT SITE (POTENTIAL).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 388 388 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 804 804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1220 1220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1317 1317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1711 1711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1733 1733 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 878 915 GEYIDDOGCHGCEACSCAMGPTDCISCPTRVLD ->
FT (IN ISOFORM PCS5A).
FT VARSPPLIC 916 1877 MISSING (IN ISOFORM PCS5A).
FT SEQUENCE 1877 AA; 209287 MW; EC850E2DF20EA1C3 CRC64;

```

Query Match 17.4%; Score 86.5; DB 1; Length 1877;
 Best Local Similarity 22.1%; Pred. No. 0.36; Mismatches 43; Indels 53; Gaps 3;
 Matches 29; Conservative 6;

```

QY 1 CSONEYFDSLHACIPQRLCSSNTPPLTCORCEYFDSL----- 42
DB 1481 CAVEVWDESHRCOPCHKSCSGSEDCYTCREFTLLMTQCKECPBGVHTDKDS 1540
QY 43 -----HACPCLRCSP-----PTCOYCFHSEYFDSLHACP----- 73
DB 1541 QOCVLCHSQRCECPHSMQCLSCRPGMFQLQCKECLQCDGYGSESTGRCEKCDKSK 1600
QY 74 -----PATCO 78
DB 1601 SCRGPRPTDQ 1611

RESULT 6
KRA3_SHEEP STANDARD; PRT; 131 AA.
AC P02441;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Keratin, high-sulfur matrix protein, IILN3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP MEDLINE=74022242; PubMed=4584026;
RA Swart L.S., Haylett T.;
RT "Studies on the high-sulphur proteins of reduced Merino wool. Amino
RL acid sequence of protein SCMB-3A3."
CC Blochm. J. 133:641-654(1973).
CC -!- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID

```

```

CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TRYPTOPHAN KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -!- MISCELLANEOUS: THE SOURCE OF THIS PROTEIN IS MERINO WOOL.
DR PIR: A02840; KRSH3.
DR InterPro: IPR002494; Keratin_B2.
DR Pfam: PF01500; Keratin_B2_1.
KW Keratin.
SQ SEQUENCE 131 AA; 14163 MW; 8C0B56C0070B1C50 CRC64;

```

Query Match 16.7%; Score 83; DB 1; Length 131;
 Best Local Similarity 25.2%; Pred. No. 0.082;
 Matches 26; Conservative 12; Mismatches 35; Indels 30; Gaps 6;

```

QY 1 CSONEYFDSLHACIPQRLCSSNTPPLT-----CQR-YCCYFDSLHAA 44
DB 18 CLQPRVYRD-----PCCRPVSCQTVSRPVTVPCTRPICERCRPVCCDPCSLQEGC 71
QY 45 CPCLRCSPPCQY-----CCFHSEYFDSLHACP---PATCOP 79
DB 72 CPPTCCPTSCQAVVCRPCCATTCQCPVSVQCPCRPRTSCOP 114

```

```

RESULT 7
UBPM_MOUSE STANDARD; PRT; 526 AA.
ID UBPM_MOUSE 061068;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase DUB-1 (EC 3.1.2.15) (Ubiquitin
DE thiolesterase DUB-1) (Ubiquitin-specific processing protease DUB-1)
DE (Deubiquitinating enzyme 1).
GN DUB1 OR DUB-1.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194957; PubMed=8622927;
RA Zhu Y., Carroll M., Papa F.R., Hochstrasser M., D'Andrea A.D.;
RT "DUB-1, a deubiquitinating enzyme with growth-suppressing activity."
RL Proc. Natl. Acad. Sci. U.S.A. 93:3275-3279(1996).
CC -!- FUNCTION: HAS GROWTH-SUPPRESSING ACTIVITY, INDUCES ARREST IN G1
CC PHASE UPON CONTROLLED EXPRESSION.
CC -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -!- INDUCTION: BY INTERLEUKIN-3.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19. ALSO KNOWN AS
CC FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U41636; AAC52532.1; -
CC MEROPS: C19.031; -
DR MGD; MGI:107699; DUB1.
DR InterPro: IPR001394; UCH-2.
DR Pfam: PF00443; UCH-1; 1.
DR PROSITE: PS00972; UCH_2_1; 1.
DR PROSITE: PS00973; UCH_2_2; 1.
DR PROSITE: PS0235; UCH_2_3; 1.
KW Ubiquitin conjugation; Hydrolase; Thiol protease; Multigene family.
FT ACT_SITE 60 60 BY SIMILARITY.

```

FT ACT_SITE 298 298 BY SIMILARITY.
DR ACT_SITE 307 307 BY SIMILARITY.
SQ SEQUENCE 526 AA; 59073 MW; 263AA7B579694FEA CRC64;
Matches 19; Conservative 8; Mismatches 18; Indels 8; Gaps 2;
Query Match 16.5%; Score 82; DB 1; Length 526;
Best Local Similarity 35.8%; Pred. No. 0.33;
Matches 19; Conservative 8; Mismatches 18; Indels 8; Gaps 2;
OY 19 LKCSSTPLTCORCYCEFFDLSLHACPLRCSPPTCCYCFHFSEYFDSLHA 71
DB 66 LQCLHTHTPL-----ADWMLSGHSGTC--CSPECKCKLAMEALVQSLHS 110
RESULT 8
LMG3_HUMAN STANDARD; PRT; 1587 AA.
AC O9Y6N6;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).
GN LAMC3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=99242614; PubMed=10225960;
RA Koch M., Olson P., Aldus A., Jin W., Hunter D.D., Brunken W.J.,
RT Burgess R.E., Champilaud M.F.;
RT "Characterization and expression of the laminin gamma3 chain: a novel,
RT non-basement membrane-associated, laminin chain.";
RL J. Cell Biol. 145:605-618(1999).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDiate THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and
CC the reproductive tracts.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF041835; AAD36991.1; -
DR HSSP; P02468; TITLE.
DR MIM; 604349; -
DR InterPro; IPR0000561; EGF-like.
DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001886; LAMNT.
DR Pfam; PF00052; laminin_B.1.
DR Pfam; PF00053; laminin_EGF.10.
DR Pfam; PF00055; laminin_Nterm.1.

DR PRINTS; PR00011; EGFLAMININ.
DR ProDom; PD002082; LAMNT.1.
DR ProDom; PD003031; Laminin_B.1.
DR SMART; SM00180; EGF_Lam.10.
DR SMART; SM00001; EGF_Like.1.
DR SMART; SM00281; Lam.1.
DR SMART; SM00136; LAMNT.1.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_Type_EGF.10.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 19
FT CHAIN 1 19
FT DOMAIN 20 1587
FT DOMAIN 20 270
FT DOMAIN 271 326
FT DOMAIN 327 382
FT DOMAIN 383 429
FT DOMAIN 430 479
FT DOMAIN 480 489
FT DOMAIN 490 672
FT DOMAIN 673 706
FT DOMAIN 707 754
FT DOMAIN 755 809
FT DOMAIN 810 865
FT DOMAIN 866 916
FT DOMAIN 917 964
FT DOMAIN 965 1013
FT DOMAIN 1014 1587
FT DOMAIN 1071 1141
FT DOMAIN 1200 1229
FT DOMAIN 1424 1504
FT DOMAIN 1535 1579
FT SITE 1059 1061
FT CARBOHD 87 87
FT CARBOHD 119 119
FT CARBOHD 295 295
FT CARBOHD 328 328
FT CARBOHD 631 631
FT CARBOHD 837 837
FT CARBOHD 980 980
FT CARBOHD 1185 1185
FT CARBOHD 1518 1518
SQ SEQUENCE 1587 AA; 172051 MW; 3C86E0B5F20319 CRC64;
Query Match 16.3%; Score 81; DB 1; Length 1587;
Best Local Similarity 31.0%; Pred. No. 1.1;
Matches 22; Conservative 8; Mismatches 31; Indels 10; Gaps 4;
OY 12 HACIPQQLRCSSNPPLTCORCYCEFFDLSLHACPLRC---SPPTCCYCFHFSEYFDS 67
DB 322 HECLPC--NCSGRSECFDR---ELFSTGSGRCHRCRDHTAGPHRCQENFYHWP 376
OY 68 LHACPRATCQ 78
DB 377 RM-PCQPCDQ 386
RESULT 9
PAC4_RAT STANDARD; PRT; 937 AA.
ID PAC4_RAT
AC Q63415;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Paired basic amino acid cleaving enzyme 4 precursor (EC 3.4.21.-)
DE (Subtilisin/kexin-like protease PAC4) (Subtilisin-like proprotein
DE convertase 4) (SPC4).
GN PACE4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

Query Match	Best Local Similarity	Score	DB 1:	Length	937:
Matchs 30;	Conservative 24.4%;	Pred. No. 0.75;	Mismatches 39;	Indels 45;	Gaps 6;

```

OY      1 CSOMYFPSLLHACIPQCLRCS-----SNTPLPTQCR--YCCSEYPSLLHACP----- 46
Db      701 CPLGFGDTAARRCRRKRCACGCTCTGRSPQTQCLSCRSGFYHHQENITVCYTLCPADLYADE 760
OY      47 ---CLRCSPPTQCYC-----CFHSEYFDSLLHACPPATQ 78
Db      761 SQRLLRCH-PSQCRCVDEPEKSTCKEGEFLARGSCIPDCBPATYFDSILIRC--GRECH 817
OY      79 PYC 81
Db      818 HTG 820

RESULT 10
ID      LMB2_HUMAN
AC      P55268; Q16321;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      01-MAR-2002 (Rel. 41, Last annotation update)
DE      Laminin beta-2 chain precursor (S-laminin) (Laminin B1s chain).
GN      LAMB2 OR LAMS.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95213013; PubMed=7698745;
RA      Wever U.M., Gerecke D.R., Durkin M.E., Kurtz K.S., Mattel M.-G.,
RA      Champiadou M.F., Burgess R.E., Albrechtsen R.;
RT      "Human beta 2 chain of laminin (formerly S chain): cDNA cloning,
RL      chromosomal localization, and expression in carcinomas.";
RN      Genomics 24:243-252(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=95316263; PubMed=7795887;
RA      Iivanainen A., Vuolteenaho R., Salnio K., Eddy R., Shows T.B.,
RA      Sariola H., Tryggvason K.;
RT      "The human laminin beta 2 chain (S-laminin): structure, expression in
RL      fetal tissues and chromosomal assignment of the LAMB2 gene.";
RN      Matrix Biol. 14:489-497(1995).
CC      -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ
CC      IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, AND ORGANIZATION OF
CC      CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC      WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC      -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC      DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC      TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC      COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC      THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ), LAMININ-4
CC      (S-MEROSIN), AND LAMININ-7 (KS-LAMININ).
CC      -1- SUBCELLULAR LOCATION: Extracellular.
CC      -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC      COMPONENT). S-LAMININ IS CONCENTRATED IN THE SYNAPTIC
CC      CLEFT OF THE NEUROMUSCULAR JUNCTION.
CC      -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC      WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC      -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC      -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC      -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC      -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/annouce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DB      EMBL; Z68155; CAA92279.1; -.

```

DR EMBL: Z68156; CA92279.1; JOINED.
 DR EMBL: X79683; CA56130.1; -.
 DR EMBL: S77512; AB34682.2; -.
 DR HSSP: P02468; IKLO.
 DR MIM: 150325; -.
 DR InterPro: IPR000561; EGF-Like.
 DR InterPro: IPR00186; Lamnt.
 DR InterPro: IPR002049; Laminin_EGF.
 DR Pfam: PF00053; Laminin_EGF_13.
 DR Pfam: PF00055; Laminin_Nterm; 1.
 DR PRINTS: PR00011; EGFLAMININ.
 DR ProDom: PD002082; Lamnt; 1.
 DR SMART: SM00180; EGF_Lam; 11.
 DR SMART: SM00136; Lamnt; 1.
 DR PROSITE: PS00022; EGF_1; 10.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01248; LAMININ_TYPE_EGF; 12.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL 1 32
 FT CHAIN 33 1798
 FT DOMAIN 33 280 LAMININ BETA-2 CHAIN.
 FT DOMAIN 283 346 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 347 409 LAMININ EGF-LIKE 1.
 FT DOMAIN 410 469 LAMININ EGF-LIKE 2.
 FT DOMAIN 470 521 LAMININ EGF-LIKE 3.
 FT DOMAIN 522 552 LAMININ EGF-LIKE 4.
 FT DOMAIN 553 781 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 783 830 LAMININ DOMAIN IV.
 FT DOMAIN 831 876 LAMININ EGF-LIKE 6.
 FT DOMAIN 877 926 LAMININ EGF-LIKE 7.
 FT DOMAIN 927 985 LAMININ EGF-LIKE 8.
 FT DOMAIN 986 1037 LAMININ EGF-LIKE 9.
 FT DOMAIN 1038 1094 LAMININ EGF-LIKE 10.
 FT DOMAIN 1095 1142 LAMININ EGF-LIKE 11.
 FT DOMAIN 1143 1189 LAMININ EGF-LIKE 12.
 FT DOMAIN 1190 1409 LAMININ EGF-LIKE 13.
 FT DOMAIN 1410 1442 DOMAIN II.
 FT DOMAIN 1443 1798 DOMAIN ALPHA.
 FT DOMAIN 1253 1319 COILED COIL (POTENTIAL).
 FT DOMAIN 1472 1526 COILED COIL (POTENTIAL).
 FT DOMAIN 1577 1790 COILED COIL (POTENTIAL).
 FT DISULFID 283 292 BY SIMILARITY.
 FT DISULFID 285 310 BY SIMILARITY.
 FT DISULFID 312 321 BY SIMILARITY.
 FT DISULFID 324 344 BY SIMILARITY.
 FT DISULFID 347 356 BY SIMILARITY.
 FT DISULFID 349 374 BY SIMILARITY.
 FT DISULFID 377 386 BY SIMILARITY.
 FT DISULFID 389 407 BY SIMILARITY.
 FT DISULFID 410 423 BY SIMILARITY.
 FT DISULFID 412 438 BY SIMILARITY.
 FT DISULFID 440 449 BY SIMILARITY.
 FT DISULFID 452 467 BY SIMILARITY.
 FT DISULFID 470 484 BY SIMILARITY.
 FT DISULFID 472 491 BY SIMILARITY.
 FT DISULFID 493 502 BY SIMILARITY.
 FT DISULFID 505 519 BY SIMILARITY.
 FT DISULFID 783 795 BY SIMILARITY.
 FT DISULFID 785 802 BY SIMILARITY.
 FT DISULFID 804 813 BY SIMILARITY.
 FT DISULFID 816 828 BY SIMILARITY.
 FT DISULFID 831 843 BY SIMILARITY.
 FT DISULFID 833 850 BY SIMILARITY.
 FT DISULFID 852 861 BY SIMILARITY.
 FT DISULFID 864 874 BY SIMILARITY.
 FT DISULFID 877 886 BY SIMILARITY.
 FT DISULFID 879 893 BY SIMILARITY.
 FT DISULFID 896 905 BY SIMILARITY.
 FT DISULFID 908 924 BY SIMILARITY.
 FT DISULFID 927 943 BY SIMILARITY.
 FT DISULFID 929 954 BY SIMILARITY.
 FT DISULFID 956 965 BY SIMILARITY.

FT DISULFID 968 983 BY SIMILARITY.
 FT DISULFID 986 1000 BY SIMILARITY.
 FT DISULFID 988 1007 BY SIMILARITY.
 FT DISULFID 1010 1019 BY SIMILARITY.
 FT DISULFID 1022 1035 BY SIMILARITY.
 FT DISULFID 1095 1107 BY SIMILARITY.
 FT DISULFID 1097 1114 BY SIMILARITY.
 FT DISULFID 1116 1125 BY SIMILARITY.
 FT DISULFID 1128 1140 BY SIMILARITY.
 FT DISULFID 1143 1155 BY SIMILARITY.
 FT DISULFID 1145 1162 BY SIMILARITY.
 FT DISULFID 1164 1173 BY SIMILARITY.
 FT DISULFID 1176 1187 BY SIMILARITY.
 FT DISULFID 1190 1190 INTERCHAIN (PROBABLE).
 FT DISULFID 1190 1193 INTERCHAIN (PROBABLE).
 FT DISULFID 1193 1193 INTERCHAIN (PROBABLE).
 FT DISULFID 1797 1797 INTERCHAIN (PROBABLE).
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 914 914 R -> G (IN REF. 2).
 FT CONFLICT 1179 1179 G -> A (IN REF. 2).
 SQ SEQUENCE 1798 AA; 196079 MW; 9355CE5B24850CB7 CRC64;

Query Match 16.2%; Score 80.5; DB 1; Length 1798;
 Best local similarity 26.9%; Pred. No. 1.3;
 Matches 29; Conservative 7; Mismatches 31; Indels 41; Gaps 8;
 Oy 1 CSQNEYFDSL-----HCIP-----COL-RCSSNPPLTCQRYCC 35
 Db 943 CHQDEYSQIVCHCRAGYTGRLCEACAFHGFDPSPGRCQLCEGSDINPMDPD--AC 1000
 Oy 36 EYFDLHLACPLRC-----SPTCOYC--CFHSEYFDLHLACPRATC 77
 Db 1001 D-----PHTGQCLRLHTEGPHCAKCPFGQAROSCHRC--TC 1040

RESULT 11
 LMA1_HUMAN
 ID LMA1_HUMAN STANDARD: PRT; 3075 AA.
 AC P25391.
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha-1 chain precursor (Laminin A chain).
 GN LAMA1 OR LAMA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91333420; PubMed=1714537;
 RA Haaparanta T., Uitto J., Ruoslahti E., Engvall E.;
 RT "Molecular cloning of the cDNA encoding human laminin A chain.";
 RL Matrix 11:151-160(1991).
 RN [2]
 RP SEQUENCE OF 1-2628 FROM N.A.
 RX MEDLINE=91264789; PubMed=2049067;
 RA Nissinen M., Vuolteenaho R., Booc-Handford R., Kallunki P.,
 RA Tryggvason K.;
 RT "Primary structure of the human laminin A chain. Limited expression
 in human tissues.";
 RL Biochem. J. 276:369-379(1991).
 RN [3]
 RP SEQUENCE OF 2397-3072 FROM N.A.
 RX MEDLINE=89280632; PubMed=2733383;
 RA Olsen D., Nagayoshi T., Fazio M., Peltonen J., Jaakkola S.,
 RA Sandorn D., Sasaki T., Kulvaneni H., Chu M.L., Deutzmann R.,

RA Timpl R., Utto B.:
 RT "Human laminin: cloning and sequence analysis of cDNAs encoding A, B1
 RT and B2 chains, and expression of the corresponding genes in human
 RT skin and cultured cells.";
 RT Lab. Invest. 60:772-782(1989).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 CC IS THOUGHT TO MEDiate THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND
 CC LAMININ-3 (S LAMININ).
 CC -1- TISSUE SPECIFICITY: Extracellular.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
 CC COMPONENT).
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- SIMILARITY: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LAMININ DOMAINS IV.
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: X58531; CAA61418.1; -
 DR PIR: S14458; S14458.
 DR HSSP: P02468; IKID.
 DR MIM: 150320; -
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001886; LAMNT.
 DR InterPro: IPR000084; Laminin_B.
 DR InterPro: IPR002069; Laminin_EGF.
 DR InterPro: IPR001791; Laminin_B; 2.
 DR Pfam: PF00052; Laminin_B; 2.
 DR Pfam: PF00053; Laminin_EGF; 15.
 DR Pfam: PF00054; Laminin_G; 5.
 DR Pfam: PF00055; Laminin_Nterm; 1.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR ProDom: PD002082; LAMNT; 1.
 DR ProDom: PD003031; Laminin_B; 2.
 DR SMART: SM00180; EGF_Lam; 14.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00281; Lamb; 2.
 DR SMART: SM00282; Lamb; 5.
 DR SMART: SM00136; LAMNT; 1.
 DR PROSITE: PS00022; EGF_1; 11.
 DR PROSITE: PS01186; EGF_2; 2.
 DR PROSITE: PS01248; LAMININ_TYPE_EGF; 15.
 DR PROSITE: PS50025; LAM_G_DOMAIN; 5.
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
 FT SIGNAL
 FT CHAIN 1
 FT DOMAIN 18 3075 LAMININ ALPHA-1 CHAIN.
 FT DOMAIN 269 326 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 270 326 LAMININ EGF-LIKE 1.
 FT DOMAIN 327 396 LAMININ EGF-LIKE 2.
 FT DOMAIN 397 453 LAMININ EGF-LIKE 3.
 FT DOMAIN 454 502 LAMININ EGF-LIKE 4.
 FT DOMAIN 503 512 LAMININ EGF-LIKE 5 (N-TERMINAL).
 FT DOMAIN 517 708 LAMININ DOMAIN IV 1 (DOMAIN IV B).
 FT DOMAIN 709 741 LAMININ EGF-LIKE 5 (C-TERMINAL).
 FT DOMAIN 742 790 LAMININ EGF-LIKE 6.
 FT DOMAIN 791 848 LAMININ EGF-LIKE 7.
 FT LAMININ EGF-LIKE 7.

FT DOMAIN 849 901 LAMININ EGF-LIKE 8.
 FT DOMAIN 902 950 LAMININ EGF-LIKE 9.
 FT DOMAIN 951 997 LAMININ EGF-LIKE 10.
 FT DOMAIN 998 1043 LAMININ EGF-LIKE 11.
 FT DOMAIN 1044 1089 LAMININ EGF-LIKE 12.
 FT DOMAIN 1090 1149 LAMININ EGF-LIKE 13.
 FT DOMAIN 1150 1159 LAMININ EGF-LIKE 14 (N-TERMINAL).
 FT DOMAIN 1160 1361 LAMININ EGF-LIKE 15.
 FT DOMAIN 1362 1402 LAMININ EGF-LIKE 16.
 FT DOMAIN 1403 1451 LAMININ EGF-LIKE 17.
 FT DOMAIN 1452 1508 LAMININ EGF-LIKE 18.
 FT DOMAIN 1509 1555 LAMININ EGF-LIKE 19.
 FT DOMAIN 1556 2116 LAMININ EGF-LIKE 20.
 FT DOMAIN 2117 2297 LAMININ EGF-LIKE 21.
 FT DOMAIN 2297 2305 LAMININ EGF-LIKE 22.
 FT DOMAIN 2305 2481 LAMININ EGF-LIKE 23.
 FT DOMAIN 2486 2673 LAMININ EGF-LIKE 24.
 FT DOMAIN 2713 2885 LAMININ EGF-LIKE 25.
 FT DOMAIN 2890 3070 LAMININ EGF-LIKE 26.
 FT DOMAIN 3070 3700 LAMININ EGF-LIKE 27.
 FT DOMAIN 3700 3960 LAMININ EGF-LIKE 28.
 FT DOMAIN 3960 4227 LAMININ EGF-LIKE 29.
 FT SITE 2524 2536 CELL ATTACHMENT SITE.
 FT DISULFID 270 279 BY SIMILARITY.
 FT DISULFID 272 290 BY SIMILARITY.
 FT DISULFID 292 301 BY SIMILARITY.
 FT DISULFID 297 305 BY SIMILARITY.
 FT DISULFID 304 324 BY SIMILARITY.
 FT DISULFID 327 336 BY SIMILARITY.
 FT DISULFID 329 361 BY SIMILARITY.
 FT DISULFID 364 373 BY SIMILARITY.
 FT DISULFID 376 394 BY SIMILARITY.
 FT DISULFID 397 409 BY SIMILARITY.
 FT DISULFID 399 427 BY SIMILARITY.
 FT DISULFID 429 438 BY SIMILARITY.
 FT DISULFID 441 451 BY SIMILARITY.
 FT DISULFID 454 467 BY SIMILARITY.
 FT DISULFID 471 471 BY SIMILARITY.
 FT DISULFID 473 482 BY SIMILARITY.
 FT DISULFID 485 500 BY SIMILARITY.
 FT DISULFID 742 751 BY SIMILARITY.
 FT DISULFID 744 757 BY SIMILARITY.
 FT DISULFID 760 769 BY SIMILARITY.
 FT DISULFID 772 788 BY SIMILARITY.
 FT DISULFID 791 806 BY SIMILARITY.
 FT DISULFID 793 816 BY SIMILARITY.
 FT DISULFID 819 828 BY SIMILARITY.
 FT DISULFID 831 846 BY SIMILARITY.
 FT DISULFID 849 863 BY SIMILARITY.
 FT DISULFID 851 870 BY SIMILARITY.
 FT DISULFID 873 882 BY SIMILARITY.
 FT DISULFID 885 899 BY SIMILARITY.
 FT DISULFID 902 914 BY SIMILARITY.
 FT DISULFID 904 921 BY SIMILARITY.
 FT DISULFID 923 932 BY SIMILARITY.
 FT DISULFID 935 948 BY SIMILARITY.
 FT DISULFID 951 963 BY SIMILARITY.
 FT DISULFID 953 969 BY SIMILARITY.
 FT DISULFID 971 980 BY SIMILARITY.
 FT DISULFID 983 995 BY SIMILARITY.
 FT DISULFID 998 1007 BY SIMILARITY.
 FT DISULFID 1000 1014 BY SIMILARITY.
 FT DISULFID 1016 1025 BY SIMILARITY.
 FT DISULFID 1028 1041 BY SIMILARITY.
 FT DISULFID 1044 1056 BY SIMILARITY.
 FT DISULFID 1046 1063 BY SIMILARITY.
 FT DISULFID 1065 1074 BY SIMILARITY.
 FT DISULFID 1077 1087 BY SIMILARITY.
 FT DISULFID 1403 1412 BY SIMILARITY.
 FT DISULFID 1405 1431 BY SIMILARITY.
 FT DISULFID 1422 1431 BY SIMILARITY.
 FT DISULFID 1434 1449 BY SIMILARITY.
 FT DISULFID 1452 1466 BY SIMILARITY.
 FT DISULFID 1454 1476 BY SIMILARITY.

FT DISULFID 1479 1488 BY SIMILARITY.
FT DISULFID 1491 1506 BY SIMILARITY.
FT DISULFID 1509 1521 BY SIMILARITY.
FT DISULFID 1511 1528 BY SIMILARITY.
FT DISULFID 1530 1539 BY SIMILARITY.
FT DISULFID 1542 1553 BY SIMILARITY.
FT DISULFID 1556 1556 INTERCHAIN (PROBABLE).
FT DISULFID 1560 1560 INTERCHAIN (PROBABLE).
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 926 926 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1407 1407 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1579 1579 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1596 1596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1678 1678 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1689 1689 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1698 1698 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1717 1717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1804 1804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1894 1894 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 16.1%; Score 80; DB 1; Length 3075;
Best Local Similarity 23.4%; Pred. No. 2.3;
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;

OY 1 CSQNFPSL--HACIPCQ-----LRCSSNTPPLNCRQYCE-- 36
DB 831 CADGYGNPTVGESECVPCDCGNDPSEBAGCDSVTGECLKCLNTGACR--CAGGF 889
OY 37 YFDSL--LHACPC-----LRCSP-PTQOYC--CFHSEYFDSLHACP 73
DB 890 YGDAVTAKNRACECHVKGSHSAYCHLETLGCLDCRKPNTGQCCDCLHGYYGLDSGHGR 949
OY 74 PATC 77
DB 950 PCNC 953

RESULT 12
YNC3_YEAST
ID_YNC3_YEAST STANDARD: PRT: 965 AA.
AC P53971;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 108.5 kDa protein in UME3-HDA1 intergenic region.
CM YN022C OR N2812.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RA Duesthoeft A., Floeth M., Filtz C., Heuss-Neitzel D.,
RA Hilbert H., Moestl D.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Andre B., Itagui Housaini I., Urrestazu L.A., Vissers S.;
RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: TO DROSOPHILA SHUTTLE CRAFT PROTEIN (STC) AND HUMAN
CC NFX1.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 271299; CAB95885.1; -
DR SGD: S0004968; YNL023C.
DR InterPro: IPR001374; R3H.
DR InterPro: IPR000967; ZNF_NFX1.
DR InterPro: IPR001841; Znf_fing.
DR Pfam: PF01424; R3H; 1.
DR Pfam: PF01422; zf-NFX-1; 5.
DR SMART: SM00393; R3H; 1.
DR SMART: SM00184; RING; 1.
DR SMART: SM00438; ZNF_NFX; 7.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NDG.
DR PROSITE: PS00589; ZF_RING_2; 1.
KW Hypothetical protein; zinc-finger; Repeat.
FT ZN_FING 68 118 RING-TYPE
FT DOMAIN 150 644 7 X APPROXIMATE REPEATS, CYS-RICH.
FT REPEAT 150 185 1.
FT REPEAT 206 251 2.
FT REPEAT 273 330 3.
FT REPEAT 352 390 4.
FT REPEAT 458 497 5.
FT REPEAT 575 610 6.
FT REPEAT 611 644 7.
SO SEQUENCE 965 AA; 108494 MW; 121C57BB07C6FA9D CRC64;

Query Match 16.0%; Score 79.5; DB 1; Length 965;
Best Local Similarity 29.6%; Pred. No. 0.96;
Matches 29; Conservative 6; Mismatches 26; Indels 37; Gaps 7;

OY 12 HACT-PCQLRCSSNTPPLNCRQYCE-----CEYFDSLHACPC-----LRCSP--P 53
DB 456 HICLPCMLTLSCGII--KQARKCPGCKPCPLSDSDLVCPGNTVPAPVRCGTLP 513
OY 54 TCQYCCF-----HSEYFDSLHACPPAT 76
DB 514 TGNHPCIKVVRGSETCGKHPMPTCH--SLDVSCPCT 549

RESULT 13
HMCT_BOMMO
ID_HMCT_BOMMO STANDARD: PRT: 3133 AA.
AC P96092;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hemocytin precursor (Humoral lectin).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-FUYOU X TOKAI; TISSUE-Hemocyte;
RX MEDLINE=95178544; PubMed=7873598;
RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Tanai K., Kado-no-Okuda K., Kato Y., Mori H.;
RT "Cloning and expression of the gene of hemocytin, an insect humoral
RT lectin which is homologous with the mammalian von Willebrand
RT factor";
RL Blochim. Biophys. Acta 1260:245-258(1995).
RN [2]
RP SEQUENCE OF 2221-3133 FROM N.A.
RA Kotani E., Iwamoto S.I., Tashiro M., Mori H., Sumida M.,
RA Matsubara F., Yamakawa M.;
RL Submitted (JAN-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR
CC ENCASCULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL
CC METAMORPHOSIS.

FT	DISULFID	2991	3054		BY SIMILARITY.
FT	DISULFID	3004	3070		BY SIMILARITY.
FT	DISULFID	3020	3072		BY SIMILARITY.
FT	DISULFID	?	3075		BY SIMILARITY.
FT	CARBOHYD	151	151	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	237	237	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	564	564	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1170	1170	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1387	1387	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1622	1622	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1727	1727	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1847	1847	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1975	1975	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	1985	1985	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2093	2093	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2113	2113	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2161	2161	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2276	2276	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2451	2451	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2647	2647	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2654	2654	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2663	2663	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2794	2794	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2810	2810	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2865	2865	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2929	2929	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	2964	2964	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	CARBOHYD	3028	3028	N-LINKED (GLCNAC. .)	(POTENTIAL).
FT	VARIANT	1288	1288	R -> G.	
FT	VARIANT	1305	1305	T -> S.	
SO	SEQUENCE	3133	AA: 343350 MW: E52100SD14A7B2B3 CRC64:		
Query Match		16.0%;	Score 79.5;	DB 1:	Length 3133;
Best Local Similarity		28.4%;	Pred. No. 2.6;		
Matches 27;		Conservative	8;	Mismatches 33;	Indels 27; Gaps
OY	12 HAC-----IPCOLRCSSTNPPLTCRCYCCCFDSLHACP-----CLRGSPPT 54		: : : : : : : :		
Db	488 HACHARGVTENMR.-TNDLCPMQCEVOSNY.-DS-CVSACPVETCDNIIYYAETTRACEODT 54549				
OY	55 C-----OVCCHSEY-FDSLHACPATGCPVC 81				
Db	546 CVEGCKPKKSCPEGSVYKNDSTTECVPRACKPCVC 580				
RESULT 14					
PCK5_BRACL	STANDARD:	PRT:	1696	AA.	
ID	PCK5_BRACL				
AC	Q9NJL5: Q9NJL5: Q9NJL4,				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Protein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)				
GN	PC6.				
OS	Branchiostoma californiensis (California lancelet) (Amphioxus).				
OC	Eumetazoa; Metazoa; Chordata; Cephalochordata; Branchiostomidae;				
OC	Branchiostoma.				
OX	NCBI_TaxID=7738:				
RN	[1]				
RP	SEQUENCE FROM N.A. (ISOFORMS A: B AND C).				
RX	MEDLINE=20175281; PubMed=10708868;				
RA	Oliva A.A., Jr., Chan S.J., Steiner D.F.:				
RT	"Evolution of the prohormone convertases: identification of a				
RL	homologue of PC6 in the protochordate amphioxus.";				
Biochim. Biophys. Acta	1477:338-348(2000).				
-I-	FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPROTEASE ACTIVITY				
CC	WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE				
CC	OF CLEAVAGE AT THE R(K/R)R CONSENSUS MOTIF (BY SIMILARITY).				
CC	-I- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR				
CC	PROTEINAINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-I-ZAA BONDS, WHERE XAA				

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:49:05 ; Search time 26.27 Seconds
(Without alignments)
533.407 Million cell updates/sec

Title: US-09-854-864-13
Perfect score: 498
Sequence: 1 CSQNEYFDLSLHACIPCOLR.....SEYFDLSLHACPPATCPQYIC 81

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organeller:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	136	27.3	185	11	088472 mus musculus
2	100.5	20.2	1792	13	057484 gallus galli
3	98	19.7	249	11	090823 mus musculus
4	98	19.7	249	11	090823 mus musculus
5	94.5	19.0	223	11	062220 mus musculus
6	94.5	19.0	223	11	064507 mus musculus
7	93.5	18.8	195	11	090141 mus musculus
8	93	18.7	293	4	014836 mus musculus
9	92.5	18.6	169	4	014836 homo sapien
10	92.5	18.6	227	11	090225 mus musculus
11	91.5	18.4	136	4	09BYR5 homo sapien
12	91.5	18.4	186	4	09BYR2 homo sapien
13	91.5	18.4	186	4	064526 mus musculus
14	91.5	18.4	191	11	0903H7 mus musculus
15	89	17.9	159	4	09BYR3 homo sapien
16	89	17.9	1574	11	088281 rattus norv

17	88	17.7	166	4	09BYR3	09byr3 homo sapien
18	87.5	17.6	154	4	09BYR2	09byr2 homo sapien
19	87.5	17.6	174	4	09BYR4	09byr4 homo sapien
20	87.5	17.6	193	4	09BYR5	09byr5 homo sapien
21	87	17.5	159	4	09BYR0	09byr0 homo sapien
22	87	17.5	165	11	09D7P3	09d7p3 mus musculus
23	87	17.5	195	11	09D0X9	09d0x9 mus musculus
24	87	17.5	1679	5	024301	024301 drosophila
25	86	17.3	188	11	070148	070148 rattus norv
26	85.5	17.2	154	4	09BYR9	09byr9 homo sapien
27	85.5	17.2	177	11	09D644	09d644 mus musculus
28	85.5	17.2	189	11	09D527	09d527 mus musculus
29	85	17.1	202	11	091W93	091w93 mus musculus
30	85	17.1	325	10	094HS1	094hs1 oryza sativ
31	84.5	17.0	767	13	09DGR2	09dgr2 xenopus lae
32	84	16.9	195	4	09BYR6	09byr6 homo sapien
33	84	16.9	201	4	09B066	09b066 homo sapien
34	84	16.9	210	4	09BYR0	09byr0 homo sapien
35	84	16.9	1671	5	09NIV5	09niv5 biophthalmari
36	83.5	16.8	110	5	09BIR2	09bir2 paramacium
37	83.5	16.8	130	11	09Z287	09z287 mus musculus
38	83.5	16.8	191	4	09BYR8	09byr8 homo sapien
39	83	16.7	175	4	007628	007628 homo sapien
40	82	16.5	98	4	09BYR8	09byr8 homo sapien
41	82	16.5	126	6	028707	028707 oryctolagus
42	82	16.5	167	11	09D122	09d122 mus musculus
43	82	16.5	168	11	09D732	09d732 mus musculus
44	81.5	16.4	85	11	070555	070555 mus musculus
45	81.5	16.4	429	13	091650	091650 xenopus lae

ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	185 AA.
088472	088472	088472		
AC	088472	088472		
DT	01-NOV-1998 (TREMUREL 08, last sequence update)			
DT	01-NOV-1998 (TREMUREL 08, last sequence update)			
DT	01-JUN-2001 (TREMUREL 17, last annotation update)			
DE	B-CELL MATURATION PROTEIN (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 17).			
DE	MEMBER 17).			
GN	TNFRSF17.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/C; TISSUE=SPLEN;			
RA	Madry C., Laabi Y., Callebaut I., Roussel J., Hatzoglu A.,			
RA	Leconiat M., Mornon J.-P., Berger R., Tsapis A.;			
RT	"Murine BCMA: a new member of the Tumor Necrosis Factor Receptor Superfamily."			
RL	Submitted (Apr-1998) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=COLON.			
RX	MEDLINE=2108560; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Kuehl J., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kaota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barz G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,			

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AF061505; AAC23799.1; -
 DR EMBL: AK020247; BAB32038.1; -
 DR MGD: MGI:1343050; T01fscf17.
 SO SEQUENCE 185 AA; 20442 MW; 8806352BAFD26A8E CRC64;

Query Match 27.3%; Score 136; DB 11; Length 185;
 Best Local Similarity 70.6%; Pred. No. 1.7e-09;
 Matches 24; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

OY 1 CSONEYFDSLHACIPQOLRCSSNPPPLTCORC 34
 DB 5 CPHESEFDSLHACKRCHLRCSN--PPATCOPYC 36

RESULT 2

OS7484 PRELIMINARY; PRT; 1792 AA.

AC 057484;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE LAMININ BETA 2-LIKE CHAIN.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-93015947; PubMed-1400373;

RA O'Rear J.J.;

RT "A novel laminin B1 chain variant in avian eye.";

RL J. Biol. Chem. 267:20555-20557(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-96209634; PubMed-9550264;

RA Liu J., Swadlow S., Xie W., Brewton R.G., Mayne R.;

RT "Primary structure and expression of a chicken laminin beta chain:
 evidence for four beta chains in birds.";

RL Matrix Biol. 16:471-481(1998).

DR EMBL: AF038555; AAB92586.1; -

DR HSBP; P02468; IKLO.

DR InterPro: IPR002106; AA_TRNA_Ligase_II.

DR InterPro: IPR000561; EGF-Like.

DR InterPro: IPR002049; Laminin_EGF.

DR InterPro: IPR001886; LamNT.

DR Pfam: PF00053; Laminin_EGF_13.

DR Pfam: PF00055; Laminin_Nterm; 1.

DR PRINTS: PR00011; EGF_LAMININ.

DR PRODOM: PD002082; LamNT; 1.

DR SMART: SM00180; EGF_Lam; 13.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DR PROSITE: PS00022; EGF_1; UNKNOW_10.

DB 1017 CPGYGGDAMRHSC-----RRSCNTLGTBPNTGPGQCCODRSGCHCLPHVSGOSCDR 1072
 OY 50 CSP-----PTCYCCFHESEFDSLHACPPATCOPYC 81
 DB 1073 CSPFNWNLGSGGCEPCACHPQH--SLSPACNOFTQCCSC 1110

RESULT 3

OS7484 PRELIMINARY; PRT; 249 AA.

AC 09DB23;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 1200009E08RIK PROTEIN.

GN TNFRSF13B OR 1200009E08RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=LUNG;

RX MEDLINE-21085660; PubMed-11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.,

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL: AK004668; BAB23457.1; -

DR MGD: MGI:189411; T01fscf13b.

SO SEQUENCE 249 AA; 26981 MW; 6F4290D719FEA037 CRC64;

Query Match 19.7%; Score 98; DB 11; Length 249;
 Best Local Similarity 25.6%; Pred. No. 0.00012;
 Matches 23; Conservative 12; Mismatches 27; Indels 28; Gaps 4;

OY 1 CSONEYFDSLHACIPQOLRCSSNPPPLTCORCCEYFDSLHACPLRCSPPTCYCCF 60
 DB 6 CPKDYWDSSRSCVSCALTCGORS-----QRTCTDE-----KCFINCRE----- 46

OY 61 HSEYFDSLHAC-----PPATCOPYC 81
 DB 47 QGRYDHLGACVSCDSTCTGHPQCAHRC 76

RESULT 4

OS7484 PRELIMINARY; PRT; 249 AA.

AC 09ER35;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE TACI PROTEIN.

GN TNFRSF13B.

OS Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=10881172;
RX MEDLINE=2117254;
RA Yan M., Marsters S.A., Grewal I.S., Wang H., Ashkenazi A., Dixit V.M.;
RT "Identification of a receptor for Blys demonstrates a crucial role in
humoral immunity";
RL Nat. Immunol. 1:37-41(2000).
DR EMBL: AF257673; AKG00081.1; -.
DR MGI: 1899411; Tntst13b.
SQ SEQUENCE 249 AA; 26947 MW; CB2F2D61C2931D81 CRC64;

Query Match
Best Local Similarity 19.7%; Score 98; DB 11; Length 249;
Matches 23; Conservative 12; Mismatches 27; Indels 28; Gaps 4;

OY 1 CSQNEFYDSLHACIPQALRCSSTNPPLTCQRYCCEYFDSLHACPCPLRCSPTQYCCF 60
DB 6 CRRDQYWDSSRKSCVSCALTCSSRS-----QRTCTDF-----CKFINCRKE----- 46
OY 61 HSEYFDSLHAC-----PPATQPYC 81
DB 47 QGRYDHLGACVSCDSTQHPQCAHFC 76

RESULT 5
O62220 PRELIMINARY; PRT; 223 AA.
AC O62220;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE SERINE 2 ULTRA HIGH SULFUR PROTEIN.
GN KRTAP5-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=91154184; PubMed=1840598;
RA Wood L., Mills M., Hatzenbuehler N., Vogeli G.;
RT "Additions and Corrections: Serine-rich ultra high sulfur protein gene
expression in murine hair and skin during the hair cycle.";
RL J. Biol. Chem. 265:4024-4024(1991).
DR EMBL: M37760; AAA40107.1; -.
DR HSSP: P01064; 1P12.
DR MGI: 1354758; Krtap5-4.
DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; UNKNOWN_2.
DR PROSITE: PS01208; VMFC; UNKNOWN_2.
SQ SEQUENCE 223 AA; 21442 MW; C654BDB9FD08C59A CRC64;

Query Match
Best Local Similarity 19.0%; Score 94.5; DB 11; Length 223;
Matches 24; Conservative 3; Mismatches 27; Indels 27; Gaps 5;

OY 1 CSQNEFYDSLHACIPQALRCSSTNPPLTCQRYCCEYFDSLHACPCPLRCSPTQYCCF 60

```

```

DB 150 CCOS-----CCKPC---CSSGCGSSCCOSSCC-----PCC-COSSCCPKPCC 189
OY 61 HSEYFDSLHACPPATQPYC 81
DB 190 QSS-----CCKPCCOSSC 203

RESULT 6
O64507 PRELIMINARY; PRT; 230 AA.
AC O64507;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE SERINE 1 ULTRA HIGH SULFUR PROTEIN.
GN KRTAP5-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=91065960; PubMed=2250030;
RA Wood L., Mills M., Hatzenbuehler N., Vogeli G.;
RT "Serine-rich ultra high sulfur protein gene expression in murine hair
and skin during the hair cycle.";
RL J. Biol. Chem. 265:21375-21380(1990).
DR EMBL: M37759; AAA40106.1; -.
DR HSSP: P01064; 1P12.
DR MGI: 1354732; Krtap5-1.
DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR001305; DnaJ_CXXCXXG.
DR InterPro: IPR001007; VMFC.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
DR PROSITE: PS01177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE: PS00637; DNAJ_CXXCXXG; UNKNOWN_1.
DR PROSITE: PS01208; VMFC; UNKNOWN_3.
SQ SEQUENCE 230 AA; 21781 MW; 6CC50B41B2137C23 CRC64;

Query Match
Best Local Similarity 19.0%; Score 94.5; DB 11; Length 230;
Matches 24; Conservative 3; Mismatches 27; Indels 27; Gaps 5;

OY 1 CSQNEFYDSLHACIPQALRCSSTNPPLTCQRYCCEYFDSLHACPCPLRCSPTQYCCF 60
DB 149 CCQSS-----CCKPC---CSSGCGSSCCOSSCC-----PCC-COSSCCPKPCC 188
OY 61 HSEYFDSLHACPPATQPYC 81
DB 189 QSS-----CCKPCCOSSC 202

RESULT 7
O9D141 PRELIMINARY; PRT; 195 AA.
AC O9D141;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE 18 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
CLONE:1110030N11, FULL INSERT SEQUENCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

```


RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.T., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Holman M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Noridone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyszewski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK020700; BAB32181.1; -
 DR HSSP: O46655; 1CJH.
 DR MGD: MGI:1925013; A030009A09Rik.
 DR InterPro: IPR002494; Keratin_B2.
 DR Pfam: PF01500; Keratin_B2.
 DR PROSITE: PS01208; VMFC; UNKNOWN_1.
 SQ SEQUENCE 227 AA; 23407 MW; 9A89B93A01E13E45 CRC64;

Query Match 18.6%; Score 92.5; DB 11; Length 227;
 Best Local Similarity 25.8%; Pred. No. 0.00054;
 Matches 23; Conservative 4; Mismatches 33; Indels 29; Gaps 4;
 QY 14 CIP--COLGSSNTPLT-----CQRCCEYFDSLHACPLCRSP 53
 DB 58 CVPSCCPSCCAPCCTTCRTPCVSVSSPCSCCSCCPSCCQSSCPACPCPCPS 117
 QY 54 TC-QYCFHSEYFDSLHACPPATCPYC 81
 DB 118 CCVPVCCFPV-----CCTPVCCKPYC 138

RESULT 11
 Q9BYR5 PRELIMINARY; PRT; 136 AA.
 AC Q9BYR5:
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE KERATIN ASSOCIATED PROTEIN 4.2.
 GN KRTAP4.2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=SCALP;
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.,
 RT "Characterization of a cluster of human high/ultrahigh keratin
 RT associated proteins on chromosome 17q12-21."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ406934; CAC27573.1; -
 DR InterPro: IPR002494; Keratin_B2.
 DR Pfam: PF01500; Keratin_B2.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 SQ SEQUENCE 136 AA; 14402 MW; 51E514863925E02E CRC64;

Query Match 18.4%; Score 91.5; DB 4; Length 136;
 Best Local Similarity 29.7%; Pred. No. 0.00045;
 Matches 27; Conservative 5; Mismatches 38; Indels 21; Gaps 5;
 QY 10 LTHACIP--COLGSSNTPLT-----CQRCCEYFDSLHAC--PCLR---C 50
 DB 17 LENCPCPSCCQTTCCRTTCRPSCCVSSCCRPCCQSCQCPICSCCQTTCCRTTC 76

QY 51 SPTCYCCFSEYFDSLHACPPATCPYC 81
 DB 77 RPSCVSSCFRPPCCQSV--CCOPTCCRPSC 105

RESULT 12
 Q9BYR2 PRELIMINARY; PRT; 186 AA.
 AC Q9BYR2:
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE KERATIN ASSOCIATED PROTEIN 4.5.
 GN KRTAP4.5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RC TISSUE=SCALP;
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.,
 RT "Characterization of a cluster of human high/ultrahigh keratin
 RT associated proteins on chromosome 17q12-21."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ406937; CAC27576.1; -
 DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001304; Ictin_C.
 DR InterPro: IPR001368; TNFR_C6.
 DR Pfam: PF01500; Keratin_B2.
 DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE: PS01208; VMFC; UNKNOWN_2.
 SQ SEQUENCE 186 AA; 19916 MW; 034D9C734304F63A CRC64;

Query Match 18.4%; Score 91.5; DB 4; Length 186;
 Best Local Similarity 30.5%; Pred. No. 0.0006;
 Matches 25; Conservative 6; Mismatches 28; Indels 23; Gaps 6;
 QY 10 LTHACIP--COLGSSNT--PPLTCQRCCEYFDSLHACPLCRSPPTC-----QYCC 59
 DB 17 LENCPCPSCCQTTCCRTTCRPSCCRPCCQ--SVCYQPTC--CHPSCCISCCRPYCC 71
 QY 60 FHSYFDSLHACPPATCPYC 81
 DB 72 ESS-----CRRPCCCRPSC 85

RESULT 13
 Q64526 PRELIMINARY; PRT; 136 AA.
 AC Q64526:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE ULTRA-HIGH SULPHUR KERATIN.
 GN KRTAP9-1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RC MEDLINE=89140394; PubMed=2465353;
 RA McNab A.R., Wood L., Theriault N., Gierman T., Vogel G.,
 RT "An ultra-high sulfur keratin gene is expressed specifically during
 RT hair growth."
 RL J. Invest. Dermatol. 92:263-266(1989).
 CC -1- SIMILARITY: TO THE PLANT THIONIN FAMILY.
 DR EMBL: M27685; AAA81560.1; -
 DR MGD: MGI:130997; Krtap9-1.

DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001010; Thionin.
 DR InterPro: IPR001368; TNFR_c6.
 DR InterPro: IPR001007; WMF.
 DR Pfam: PF01500; Keratin_B2; 1.
 DR PROSITE: PS00271; THIONIN; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 DR PROSITE: PS01208; WMF; UNKNOWN_1.
 SO SEQUENCE 186 AA; 19658 MW; 61D6BFDAB72CFEB4 CRC64;

Query Match 18.4%; Score 91.5; DB 11; Length 186;
 Best Local Similarity 34.7%; Pred. No. 0.0006;
 Matches 25; Conservative 4; Mismatches 24; Indels 19; Gaps 6;

OY 14 CIP-COLRC--SSNTPPLTCORCCCFYFDSLHAC--PCLRCSPPTCCQYCFHSEYFDSL 69
 DB 85 CQPCCPSCCCSCCQPCRCSSCCQPCRCISSCCQPC--CRPSCCQSSC----- 132
 OY 70 HACPPATCQPC 81
 DB 133 --CRPC-CQPC 141

RESULT 14
 O9D3H7 PRELIMINARY; PRT; 191 AA.
 AC O9D3H7;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 01 DAYS NEONATE HEAD CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
 DE CLONE:5530401L02, FULL INSERT SEQUENCE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itochi Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aitawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombereis P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL: AK017437; BAB30743.1.
 DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001010; Thionin.
 DR Pfam: PF01500; Keratin_B2; 1.
 DR PROSITE: PS00271; THIONIN; UNKNOWN_1.
 SO SEQUENCE 191 AA; 20088 MW; 522B841DC9A8A9D5 CRC64;

Query Match 18.4%; Score 91.5; DB 11; Length 191;
 Best Local Similarity 32.9%; Pred. No. 0.00062;
 Matches 23; Conservative 3; Mismatches 21; Indels 23; Gaps 5;

OY 16 PQLRCSSNTP---PLTCORCCCFYFDSLHACPCLRCSPPTCCQYCFHSEYFDSLHA 71
 DB 63 PCCVSSCCRTCCQPCPCCVSSCCQ-----PC--CQPCSCQSSCCQPC-----C 103
 OY 72 CPPATCQPC 81
 DB 104 CQPCSCQPC 113

RESULT 15
 O9BYO3 PRELIMINARY; PRT; 159 AA.
 AC O9BYO3;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE KERATIN ASSOCIATED PROTEIN 9.3.
 GN KRTAP9.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SCALP;
 RA Rogers M.A., Langbein L., Winter H., Ehmman C., Korn B., Schweizer J.,
 RT "Characterization of a cluster of human high/ ultrahigh keratin
 RT associated proteins on chromosome 17q12-21.";
 RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ406947; CAC27586.1.
 DR InterPro: IPR002494; Keratin_B2.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF01500; Keratin_B2; 1.
 DR PROSITE: PS00652; TNFR_NGFR_1; UNKNOWN_1.
 SO SEQUENCE 159 AA; 16853 MW; 375C1E52EECD68 CRC64;

Query Match 17.9%; Score 89; DB 4; Length 159;
 Best Local Similarity 29.5%; Pred. No. 0.0011;
 Matches 26; Conservative 4; Mismatches 20; Indels 38; Gaps 7;

OY 10 LHAACP-COLRC-----SSNTPPLTCORCCCFYFDSLHAC--PCLRCSP 53
 DB 1 MHCSPCCQPCPCRTCCQPTVTTCSSTP--CQPCSC-----VSSCCQPC--CHRT 50
 OY 54 TCQYCCFHSEYFDSLHACPATCQPC 81
 DB 51 CCQNTC-----CRTCCQPC 66

Search completed: May 15, 2002, 09:53:06
 Job time: 241 sec

FT		/note= "Derived from human BCMA"
XX		
PN	WO200124811-A1.	
PD	12-APR-2001.	
XX		
PE	05-OCT-2000; 2000WO-US27579.	
XX		
PR	06-OCT-1999; 99US-0157933.	
PR	11-FEB-2000; 2000US-0181807.	
XX	30-JUN-2000; 2000US-0215688.	
PA	(BIOT) BIOGEN INC.	
XX	(APOT-) APOTECH R & D SA.	
PI	Schneider P, Thompson J, Cachero T, Ambrose C, Rennett P;	
DR	WP1; 2001-266242/27.	
DR	N-PSDB; AAD03847.	
PT	Treating a mammal for a condition associated with undesired cell	
PT	proliferation such as cancer or carcinoma, comprises administering a	
PT	composition comprising A proliferation inducing Ligand Receptor	
PT	(APRIL-R) antagonist -	
XX		
PS	Example 1; Fig 3B; 85pp; English.	
XX		
CC	The invention relates to a method of treating a mammal for a condition	
CC	associated with undesired cell proliferation such as cancer or	
CC	carcinoma. The method involves administering a composition comprising	
CC	A proliferation inducing Ligand Receptor (APRIL-R) also referred as	
CC	B cell maturation protein (BCM or BCMA) antagonist that antagonises the	
CC	interaction between APRIL and its cognate receptor(s). This method is	
CC	useful for treating undesired cell proliferation such as cancer or	
CC	carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,	
CC	prostate carcinoma, and other carcinomas whose proliferation is modulated	
CC	by APRIL. It is also useful for treating autoimmune diseases (Grave's	
CC	disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular	
CC	diseases, renal disorders, B-cell lympho-proliferative disorders,	
CC	immunosuppressive diseases, organ transplantation, inflammation and	
CC	human immunodeficiency virus (HIV), and for treating, suppressing or	
CC	altering an immune response involving a signalling pathway between	
CC	APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.	
CC	The present sequence is a fusion construct containing human APRIL-R	
CC	also referred as BCMa or BCM protein, FC region of human Immunoglobulin	
CC	G (IgG) and a signal sequence from murine Ig kappa cDNA.	
XX		
SO	Sequence 302 AA;	
	Query Match 41.2%; Score 205; DB 22; Length 302;	
	Best Local Similarity 67.8%; Pred. No. 1.9e-10;	
	Matches 40; Conservative 2; Mismatches 9; Indels 8; Gaps 3.	
OY	1 CSQNEFDSLHACIPCLRCSSNPPLTQCRR-C-CEVFDSL-----LHACPCLRCSP 53	
	:	
Dd	31 csqneyfslshacipclrcssnptltcqrqycnasvtnskvydkhtcp--pcpap 87	
RESULT 2	.	
ID	AAB60699	
XX	AAB60699 standard; Protein; 302 AA.	
AC	AAB60699;	
DT		
DT	22-MAY-2001 (first entry)	
DE		
XX	Mouse Igg signal/human BAFF-R/human Igg Fc fusion protein, BAFF-R-Fc.	
KW	Human BAFF-R; BAFF receptor; TNF family; immunoregulatory agent;	
KW	immune-related disorder; B-cell growth inhibitor;	
XX	B-cell maturation inhibitor; immunoglobulin production inhibitor;	
XX	autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;	
KW	autocrine disorder; B-cell lymphoproliferative disorder; hypertension;	

KW	renal disorder; immunosuppressive disorder; HIV infection;
KM	organ transplantation; antiinflammatory; systemic lupus erythematosus;
KV	autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
KW	B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
KM	lymphoma; gene therapy; cancer; tumour; IgG Fc; fusion construct.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Mus sp.
PX	WO200112812-A2.
PN	
PD	22-FEB-2001.
XX	
PF	16-AUG-2000; 2000MO-US22507.
XX	
PR	17-AUG-1999; 99US-0149378.
PR	11-FEB-2000; 2000US-0181684.
XX	
PR	18-FEB-2000; 2000US-0183536.
PA	(BIOJ) BIOGEN INC.
PA	(APOT) APOTEC R & D SA.
XX	
PI	Mackay F, Browning J, Ambrose C, Tschopp J, Schneider P;
PT	Thompson J;
DR	WPI, 2001-202866/20.
XX	N-PSDB; AAF59999.
XX	
PT	Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
PT	lympho-proliferative disorder by administering BAFf-R receptor
PT	polypeptide, chimeric molecule comprising receptor or anti-BAFf-R
PT	antibody homolog -
XX	
PS	Example 4; Fig 2; 59pp; English.
XX	
CC	The invention relates to the use of a BAFf receptor (BAFf-R, also known
CC	as BCMa) protein, or a BAFf-R fusion protein as an agent for the
CC	treatment of a variety of immune-related disorders. BAFf-R is a member of
CC	the TNF (tumour necrosis factor) family, acting as an immunoregulatory
CC	agent, and also plays a role in the development of hypertension and
CC	related disorders. BAFf-R, fusion proteins containing it, and BAFf-R-
CC	specific antibodies can be used for inhibiting B-cell growth, dendritic
CC	cell-induced B-cell growth and maturation, and immunoglobulin production,
CC	and in the treatment of autoimmune disorders, B-cell lymphoproliferative
CC	disorders, hypertension and renal disorders. The BAFf-R proteins may also
CC	be used in the treatment of immunosuppressive disorders and HIV
CC	infection, and in patients undergoing organ transplantation. The BAFf-R
CC	proteins or BAFf-R-specific antibodies may be used for treating,
CC	suppressing or altering an immune response involving a signalling pathway
CC	between BAFf-R and BAFf, thereby inhibiting inflammation. Since BAFf-R
CC	inhibits B-cell growth and maturation it is useful for treating diseases
CC	such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
CC	Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
CC	progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
CC	human BAFf-R may be used in gene therapy to treat tumours, lymphomas,
CC	autoimmune disorders and inherited B-cell-associated disorders. The
CC	present sequence represents the BAFf-R fusion protein BAFf-R-Fc,
CC	comprising a mouse IgG-kappa signal sequence, residues 1-153
CC	of human BAFf-R and a human IgG Fc sequence.
CC	
XX	
SQ	Sequence 302 AA:
Query Match	41.2%; Score 205; DB 22; Length 302;
Best Local Similarity	67.8%; Pred. No. 1.9e-10;
Matches 40;	Conservative 2; Mismatches 9; Indels 8; Gaps 3;
OY	1 CSQNTFYSLHACIPCOLRCSNNPPLTLCQRYC-CXEYFDLSL-----LHACPCLRCSP 53
DB :
	db 31 csqneyfdslhacipcqrlrcsnncplptlcqrycnastrvskvydkhtcp--pcpap 87

[illegible]

AC	AAY94001;
XX	
DT	20-OCT-2000 (first entry)
XX	
DE	A human BCMA protein, a B cell protein related to TACI.
XX	
XX	
KM	Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
KM	transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
KM	zntf4 activity; antibody production; autoimmune disease; amyloidosis;
KW	systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
KW	rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
KW	end stage renal failure; glomerulonephritis; vasculitis; nephritis;
KW	renal neoplasms; multiple myeloma; lymphoma; light chain neuropathy;
KM	immune response; immunosuppression; graft rejection; joint pain;
KM	graft versus host disease; inflammation; swelling; anemia; septic shock;
KW	insulin dependent diabetes mellitus; Crohn's disease; hypertension;
KW	renal artery stenosis; occlusion; cholesterol; renal emboli.
OS	
XX	Homo sapiens.
PN	WO200040716-A2.
XX	
PD	13-JUL-2000.
XX	
PF	07-JAN-2000; 2000WO-US00396.
XX	
PR	07-JAN-1999; 99US-0226533.
PA	(ZIMO) ZIMOGENETICS INC.
XX	
PI	Gross JA, Xu W, Madden K, Yee DP;
XX	
DR	WPI; 2000-452538/39.
XX	
N-PSDB:	AAA58559.
PT	Inhibiting zntf4 activity in a mammal, to treat autoimmune diseases,
PT	renal disease, graft versus host disease, and inflammation, comprises
XX	administering a BR43x2, TACI or BCMA extracellular domain polypeptide -
PS	Disclosure: Page 152; 175pp; English.
XX	
CC	The present sequence represents a human BCMA protein, a B cell protein
CC	related to transmembrane activator and CAML-interactor (TACI) receptor.
CC	TACI is a tumour necrosis factor (TNF) receptor. The extracellular
CC	domains of BR43x2 (an isoform of TACI), TACI or BCMA (a related B cell
CC	protein) receptor contain a cysteine rich domain, and are used for
CC	inhibiting zntf4 activity. Zntf4 is a TNF ligand. They may also be used
CC	for inhibiting BR43x2, TACI or BCMA receptor-ligand engagement associated
CC	with activated or resting B lymphocytes, effector T-cells, or with
CC	antibody production. The antibody production is associated with an
CC	autoimmune disease selected from systemic lupus erythematosus, myasthenia
CC	gravis, multiple sclerosis and rheumatoid arthritis. The zntf4 activity
CC	and BR43x2, TACI or BCMA receptor-ligand engagement is associated with
CC	asthma, bronchitis, emphysema, end stage renal failure,
CC	glomerulonephritis, vasculitis, nephritis, pyelonephritis, renal
CC	neoplasms, multiple myelomas, lymphomas, light chain neuropathy
CC	amyloidosis, moderating immune response, immunosuppression, graft
CC	rejection, graft versus host disease, inflammation, insulin dependent
CC	diabetes mellitus, Crohn's disease, joint pain, swelling, anemia, or
CC	septic shock. BR43x2, TACI, and BCMA polypeptides, fusions, antibodies,
CC	agonists or antagonists can be used to treat hypertension, renal artery
CC	stenosis, or occlusion, and cholesterol or renal emboli.
XX	
SQ	Sequence 184 AA:
Query Match	40.4%; Score 201; DB 21; Length 184;
Best Local Similarity	100.0%; Pred. No. 2,8e-10;
Matches 34:	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CSONEYPDSLHACIPQCRLRCSSNTPPLTCQRVC 34 8 csneyfdslhacipqcrlrcssnnppltcqrv 41

```
RESULT 5
AAE09241
ID AAE09241 standard; Protein: 184 AA.
XX
XX AAE09241:
AC
XX
XX 19-NOV-2001 (first entry)
DT
XX
DE Human BCMA protein.
XX
XX Human; TNF; tumour necrosis factor; TALL-1; APRIL; TNF receptor;
KW TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
KW autoimmune disease; rheumatoid arthritis; multiple sclerosis;
KW psoriasis.
XX
XX Homo sapiens.
OS
XX WO200160397-A1.
PN
XX 23-AUG-2001.
PD
XX
XX 28-NOV-2000; 2000WO-US32378.
PE
XX 16-FEB-2000; 2000US-0182938.
PR
XX 22-AUG-2000; 2000US-0226986.
PR
XX
XX (GETH ) GENENTECH INC.
PA
XX
XX Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Masters SA, Pitti RM;
PI Yan M;
PI
XX WPI: 2001-541628/60.
DR
XX N-PSDB; AAD15902.
DR
XX
XX Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
PT activity, for treating autoimmune disorders and cancer, comprises
PT exposing the cells to TALL-1 or APRIL polypeptide agonists or
PT antagonists.
XX
XX Example 2; Fig 2; 160pp; English.
PS
XX
XX The invention relates to methods of using one or more agonists or
CC antagonists to modulate the activity of the members of TNF (tumour
CC necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
CC e.g. TACI or BCMA. The method is useful for treating pathological
CC conditions or diseases associated with increased TALL-1 and APRIL
CC expression or activity. TALL-1 and APRIL antagonists are used to
CC block the interaction between APRIL and TALL-1 with TACI or BCMA.
CC They are useful for treating a mammal suffering from cancer such
CC as leukaemia, lymphoma, myeloma, cancers of lung and colon and
CC autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
CC psoriasis and lupus erythematosus. The present sequence is human
CC BCMA protein.
XX
XX Sequence 184 AA;
SQ
XX
XX Query Match 40.4%; Score 201; DB 22; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
XX
XX 31-JUL-2001 (first entry)
DT
XX
XX Human B cell maturation protein (BCMA).
DE
XX
XX Human; A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;
KW gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;
KW carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;
KW systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;
KW B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;
KW organ transplantation; HIV; human immunodeficiency virus; TNF;
KW tumour necrosis factor; BCMA; B cell maturation protein.
XX
XX Homo sapiens.
OS
XX WO200124811-A1.
PN
XX 12-APR-2001.
PD
XX
XX 05-OCT-2000; 2000WO-US27579.
PE
XX
XX 06-OCT-1999; 99US-0157933.
PR 11-FEB-2000; 2000US-0181807.
PR 30-JUN-2000; 2000US-0215688.
PR
XX
XX (BIOJ ) BIOGEN INC.
PA (ABOT-) APOTTECH R & D SA.
XX
XX Schneider P, Thompson J, Cachero T, Ambrose C, Rennert P;
PI
XX WPI: 2001-266242/27.
DR
XX N-PSDB; AAD03844.
DR
XX
XX Treating a mammal for a condition associated with undesired cell
PT proliferation such as cancer or carcinoma, comprises administering a
PT composition comprising A Proliferation Inducing Ligand Receptor
PT (APRIL-R) antagonist.
XX
XX Claim 3; Fig 3A; 85pp; English.
PS
XX
XX The invention relates to a method of treating a mammal for a condition
CC associated with undesired cell proliferation such as cancer or
CC carcinoma. The method involves administering a composition comprising
CC A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as
CC B cell maturation protein (BCM or BCMA) antagonist that antagonises the
CC interaction between APRIL and its cognate receptor(s). This method is
CC useful for treating undesired cell proliferation such as cancer or
CC carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,
CC prostate carcinoma, and other carcinomas whose proliferation is modulated
CC by APRIL. It is also useful for treating autoimmune diseases (Grave's
CC disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular
CC diseases, renal disorders, B-cell lympho-proliferative disorders,
CC immunosuppressive diseases, organ transplantation, inflammation and
CC human immunodeficiency virus (HIV), and for treating, suppressing or
CC altering an immune response involving a signalling pathway between
CC APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.
CC
CC The present sequence is human APRIL-R also referred as BCMA or
CC BCM protein.
XX
XX Sequence 184 AA;
SQ
XX
XX Query Match 40.4%; Score 201; DB 22; Length 184;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


ID AAB60698 standard; Protein; 184 AA.
 XX AAB60698;
 AC
 XX 22-MAY-2001 (first entry)
 DT
 XX
 DE Human BAFR receptor (BAFR-R).
 XX
 KW Human BAFR-R; BAFR receptor; TNF family; immunoregulatory agent;
 KW immune-related disorder; B-cell growth inhibitor; BCMA; B-cell
 KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
 KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
 KW renal disorder; immunosuppressive disorder; HIV infection;
 KW organ transplantation; anti-inflammatory; systemic lupus erythematosus;
 KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
 KW B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
 KW lymphoma; gene therapy; cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 PN WO200112812-A2.
 XX
 PD 22-FEB-2001.
 XX
 PF 16-AUG-2000; 2000WO-US22507.
 XX
 PR 17-AUG-1999; 99US-0149378.
 PR 11-FEB-2000; 2000US-0181684.
 PR 18-FEB-2000; 2000US-0183536.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (APOT-) APOTEC R & D SA.
 XX
 PI Mackay F, Brownlind J, Ambrose C, Tschopp J, Schneider P,
 PI Thompson J;
 DR WPI: 2001-202866/20.
 DR N-PSDB: AAF59998.
 XX
 PT Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
 PT lympho-proliferative disorder by administering BAFR-receptor
 PT polypeptide, chimeric molecule comprising receptor or anti-BAFR-R
 PT antibody homolog
 XX
 PS Claim 20; Fig 1; 59pp; English.
 XX
 CC The invention relates to the use of a BAFR receptor (BAFR-R, also known
 CC as BCMA) protein, or a BAFR-R fusion protein as an agent for the
 CC treatment of a variety of immune-related disorders. BAFR-R is a member of
 CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
 CC agent, and also plays a role in the development of hypertension and
 CC related disorders. BAFR-R, fusion proteins containing it, and BAFR-R-
 CC specific antibodies can be used for inhibiting B-cell growth, dendritic
 CC cell-induced B-cell growth and maturation, and immunoglobulin production,
 CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
 CC disorders, hypertension and renal disorders. The BAFR-R proteins may also
 CC be used in the treatment of immunosuppressive disorders and HIV
 CC infection, and in patients undergoing organ transplantation. The BAFR-R
 CC proteins or BAFR-R specific antibodies may be used for treating,
 CC suppressing or altering an immune response involving a signalling pathway
 CC between BAFR-R and BAFR, thereby inhibiting inflammation. Since BAFR-R
 CC inhibits B-cell growth and maturation it is useful for treating diseases
 CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
 CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
 CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
 CC human BAFR-R may be used in gene therapy to treat tumours, lymphomas,
 CC autoimmune disorders and inherited B-cell-associated disorders. The
 CC present sequence represents human BAFR-R.
 XX
 SO Sequence 184 AA;

Query Match

40.4%; Score 201; DB 22; Length 184;

Best Local Similarity 100.0%; Pred. No. 2.8e-10;
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSQNEIFDSLHACIPQQLRCSSNTPEPTCQRYC 34
 DB 8 csqneyfdsllhacipqqlrcssntptlctqryc 41

RESULT 8

AA71979 standard; Protein; 184 AA.

AA71979;

28-MAR-2001 (first entry)

Human B cell maturation factor (BCMA) protein.

Human; Tumour Necrosis Factor; TNF; immunosuppressant; TALL-1;
 Tumour necrosis factor and Apol-related leucocyte expressed ligand 1;
 therapy; autoimmune disorder; rheumatoid arthritis; multiple sclerosis;
 systemic lupus erythematosus; SLE; insulin dependent diabetes mellitus;
 thrombocytopenia purpura; acute rheumatic fever; Goodpasture's syndrome;
 haemolytic anaemia; Grave's disease; myasthenia gravis; chromosome 16;
 post-streptococcal glomerulonephritis; polyarteritis nodosa; BCMA;
 B cell maturation factor; pemphigus vulgaris; B-lymphocyte proliferation.

Homo sapiens.

Key Location/Qualifiers

FT Domain 1..62 /label= Extracellular_domain

WO200068378-A1.

16-NOV-2000.

05-MAY-2000; 2000WO-US12266.

06-MAY-1999; 99US-0132892.

01-MAY-2000; 2000US-0201012.

(NAJE-) NAT JEWISH MEDICAL & RES CENT.

Shu HS;

WPI: 2001-016094/02.

N-PSDB: AAD02125.

Isolated TALL-1 protein is used to identify compounds that regulate B

lymphocyte proliferation, used to treat B lymphocyte associated

autoimmune disorders -

Claim 37; Page 104-105; 112pp; English.

The present invention relates to tumour necrosis factor (TNF) and
 Apol-related leucocyte expressed ligand 1 (TALL-1) nucleic acid
 molecules, proteins (including homologues), and their antibodies. The
 invention in particular relates to methods for regulating the
 interaction between TALL-1 and TALL-1 receptors (BCMA referred as B cell
 maturation factor) to regulate monocyte, macrophage and B lymphocyte
 mediated immune responses. TALL-1 protein is useful for identifying
 compounds that regulate B lymphocyte proliferation. It is also useful for
 treating B lymphocyte associated autoimmune disorders like rheumatoid
 arthritis, systemic lupus erythematosus (SLE), insulin dependent diabetes
 mellitus, multiple sclerosis, myasthenia gravis, Grave's disease,
 autoimmune haemolytic anaemia, autoimmune thrombocytopenia purpura,
 Goodpasture's syndrome, pemphigus vulgaris, acute rheumatic fever,
 post-streptococcal glomerulonephritis, or polyarteritis nodosa.
 The TALL-1 protein and its corresponding nucleic acid sequence are also
 useful in diagnostic assays.
 The present sequence is a human B cell maturation factor (BCMA)
 protein. It is the receptor for TALL-1 protein. BCMA gene is

located on chromosome 16. In human tissues, BCMA is expressed by spleen and lymph nodes but not by brain, muscle, heart, lung, kidney, pancreas, testis and placenta. BCMA mRNA is absent in the pro-B lymphocyte stage but its expression increases with B lymphocyte maturation.

Sequence 184 AA:

Query Match 40.4%; Score 201; DB 22; Length 184;
Best Local Similarity 100.0%; Pred. No. 2,8e-10;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CSQNEYPDSLHACIPQQLRCSSNTPPLTCQRYC 34
|||||
Db 8 csqneyfdslhacipqqlrcssnptlctqryc 41

RESULT 9
AAB60700
ID AAB60700 standard; Protein: 157 AA.
XX
AC AAB60700;
XX
DT 22-MAY-2001 (first entry)
XX
DE Human BAF-R receptor (BAF-R) sequence encoded by A plasmid pJST535.
XX
KW Human BAF-R; BAF-R receptor; TNF family; immunoregulatory agent;
KW B-cell maturation inhibitor; immunoglobulin production inhibitor;
KW autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;
KW renal disorder; immunosuppressive disorder; HIV infection;
KW organ transplantation; antiinflammatory; systemic lupus erythematosus;
KW autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
KW B-cell carcinoma; leukemia; rapidly progressive glomerulonephritis;
KW lymphoma; gene therapy; cancer; tumour; plasmid pJST535.
XX
OS Homo sapiens.
XX
PN WO200112812-A2.
XX
PD 22-FEB-2001.
XX
PF 16-AUG-2000; 2000WO-US22507.
XX
PR 17-AUG-1999; 99US-0149378.
PR 11-FEB-2000; 2000US-0181684.
PR 18-FEB-2000; 2000US-0183536.
XX
PA (BIOJ) BIOGEN INC.
PA (APOT-) APOTEC R & D SA.
XX
PI Mackay F, Browning J, Ambrose C, Tschopp J, Schneider P;
PI Thompson J;
XX
DR WPI; 2001-202866/20.
DR N-PSDB: AAF60000.
XX
PT Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
PT lympho-proliferative disorder by administering BAF-R-receptor
PT polypeptide, chimeric molecule comprising receptor or anti-BAF-R
PT antibody homolog -
XX
PS Example 1; Fig 3; 59pp; English.
XX
CC The invention relates to the use of a BAF-R receptor (BAF-R, also known
CC as BCMA) protein, or a BAF-R fusion protein as an agent for the
CC treatment of a variety of immune-related disorders. BAF-R is a member of
CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory
CC agent, and also plays a role in the development of hypertension and
CC related disorders. BAF-R, fusion proteins containing it, and BAF-R-
CC specific antibodies can be used for inhibiting B-cell growth, dendritic

cell-induced B-cell growth and maturation, and immunoglobulin production,
CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative
CC disorders, hypertension and renal disorders. The BAF-R proteins may also
CC be used in the treatment of immunosuppressive disorders and HIV
CC infection, and in patients undergoing organ transplantation. The BAF-R
CC proteins or BAF-R specific antibodies may be used for treating,
CC suppressing or altering an immune response involving a signalling pathway
CC between BAF-R and BAF, thereby inhibiting inflammation. Since BAF-R
CC inhibits B-cell growth and maturation it is useful for treating diseases
CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
CC Grave's disease, multiple myeloma, B-cell carcinomas, leukemia, rapidly
CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
CC human BAF-R may be used in gene therapy to treat tumours, lymphomas,
CC autoimmune disorders and inherited B-cell-associated disorders. The
CC present sequence represents a human BAF-R protein sequence as encoded
CC by plasmid pJST535. However, this BAF-R protein sequence is 27 amino
CC acids shorter than that given in AAB60698.
XX
SQ Sequence 157 AA:

Query Match 32.0%; Score 159.5; DB 22; Length 157;
Best Local Similarity 90.6%; Pred. No. 9.1e-07;
Matches 29; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

OY 3 QNEYFDSLHACIPQQLRCSSNTPPLTCQRYC 34
|||||
Db 7 qneyfdslhacipqqlr---ntplctqryc 35

RESULT 10
AAB08844
ID AAB08844 standard; peptide: 185 AA.
XX
AC AAB08844;
XX
DT 02-JAN-2001 (first entry)
XX
DE Amino acid sequence of murine BCMA polypeptide.
XX
KW BCMA; necrosis factor-KB activator; NF-KB; gene expression; cancer;
KW anti-cell death gene; apoptosis; viral infection; inflammatory response;
KW rheumatoid arthritis; inflammatory bowel disease; septic shock.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Domain 47..72
FT /note= "putative transmembrane domain"
XX
PN WO200050633-A1.
XX
PD 31-AUG-2000.
XX
PF 24-FEB-2000; 2000WO-US04925.
XX
PR 24-FEB-1999; 99US-0121485.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Seed B, Ting A.
XX
DR WPI; 2000-558405/51.
XX
PT Identifying a modulator of gene expression for drug designing, by
PT contacting a compound library with a cell expressing an anti-cell death
PT gene and reporter gene, and determining alteration in reporter gene
PT expression -
XX
PS Claim 32; Fig 7B; 53pp; English.
XX
CC The present sequence represents a BCMA (not defined) polypeptide. BCMA
CC is a necrosis factor (NF)-KB activator. The method of the invention is

CC used to identify compounds which modulate BCMA activity (and thus NF- κ B activity). The specification describes a method of identifying a polypeptide which increases gene expression from a promoter. The method involves contacting a library of with a cell which expresses a recombinant anticell death gene and a reporter gene operably linked to the promoter, and then determining whether the expression of the reporter gene is altered as a result of contact with library. The method is useful for identifying polypeptides which increase or decrease gene expression from a promoter. The BCMA polypeptide or nucleic acid are useful for preparing a pharmaceutical composition for treating cancer, apoptosis, viral infections, inflammatory response, such as rheumatoid arthritis, inflammatory bowel disease or septic shock. BCMA is useful for identifying compounds that modulate NF- κ B expression and thus for drug designing.

Sequence 185 AA

Query Match	27.3%	Score 136;	DB 21;	Length 185;
Best Local Similarity	70.6%	Pred. No. 0.00011;		
Matches 24; Conservative	2;	Mismatches 6;	Indels 2;	Gaps 1.

```

QY      1 CSQNEYFDLSLHACIPCOLRCSSTNPPLTCORYC 34
        | :||||||| | | ||| : | | | | |
Db      5 cfhseyfdslhacpchrscn--ppatcpgyc 36

```

RESULT

ID	AAV71980 standard	Protein; 185 AA
----	-------------------	-----------------

AC AAY71980;

DT 28-MAR-2001 (first entry)

Murine B cell maturation factor (BCMA) protein.

KM	Tumour necrosis factor- α -related leucocyte expressed ligand 1;
KM	tumour necrosis factor and ApoJ-related leucocyte expressed ligand 1;
KM	therapy; autoimmune disorder; rheumatoid arthritis; multiple sclerosis;
KM	systemic lupus erythematosus; SLE; insulin dependent diabetes mellitus;
KM	thrombocytopenia purpura; acute rheumatic fever; Goodpasture's syndrome;
KM	haemolytic anaemia; Grave's disease; myasthenia gravis; BCMA;
KM	B cell maturation factor; pemphigus vulgaris; B lymphocyte proliferation;
KM	post-streptococcal glomerulonephritis; polyarthritis nodosa.

05 Mus musculus

WO200068378-A1.

PD 16-NOV-2000.

05-MAY-2000; 2000WD-US12266.

PR 06-MAY-1999; 99US-0132892.

XX
XX

PA (NAJE-) NAT JEWISH MEDICAL & RES CENT.

PI Shu HS;

DR WPI; 2001-016094/02.

DR N-PSDB; AAD02130

PT isolated TALL-1 protein is used to identify compounds that regulate B lymphocyte proliferation, used to treat B lymphocyte associated autoimmune diseases

PS Claim 37; Page 107-108; 112pp; English.

CC The present invention relates to Tumour necrosis factor (TNF) and
CC Apol-related leucocyte expressed Ligand 1 (TALL-1) nucleic acid
CC molecules, proteins (including homologues), and their antibodies. The

CC inventing in particular relates to methods for regulating the
CC interaction between TALL-1 and TALL-1 receptors (BCMA referred as B cell
CC maturation factor) to regulate monocyte, macrophage and B lymphocyte
CC mediated immune responses. TALL-1 protein is useful for identifying
CC compounds that regulate B lymphocyte proliferation. It is also useful for
CC treating B lymphocyte associated autoimmune disorders like rheumatoid
CC arthritis, systemic lupus erythematosus (SLE), insulin dependent diabetes,
CC mellitus, multiple sclerosis, myasthenia gravis, Grave's disease,
CC autoimmune haemolytic anaemia, autoimmune thrombocytopenia purpura,
CC Goodpasture's syndrome, pemphigus vulgaris, acute rheumatic fever,
CC post-streptococcal glomerulonephritis, or polyarthritis nodosa.
CC The TALL-1 protein and its corresponding nucleic acid sequence are also
CC useful in diagnostic assays.
CC The present sequence is a murine B cell maturation factor (BCMA).
CC BCMA is the receptor for TALL-1 protein.
CC

BCMA is the receptor for TALL-1 protein.

50 Sequence 185 AA;

50 Sequence 185 AA;

Query Match	27.3%	Score 136; DB 22;	Length 185;
Best Local Similarity	70.6%	Pred. No. 0.00011;	
Matches 24; Conservative	2;	Mismatches 6;	Indels 2; Gaps 1;

```

QY      1 CSQNEYFDSLIHACIPQCLRCSSNTPPLTCQRYC 34
        | : ||||| || ||||| || ||||| ||
Db      5 cfhseyfdslhacpchlrcsn--pqtacqpyc 36

```

RESULT 12

ID AAY94006 standard; Protein; 249 AA

AC AAY 94006;

DT 20-OCT-2000 (first entry)

DE A murine ztnf4, a tumour necrosis factor ligand

KM Human: E43K32: TACI receptor; extracellular domain; BCMA; B cell protein;
KM transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
KM znuA activity; antibody production; autoimmune disease; amyloidosis;
KM systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
KM rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
KM end stage renal failure; glomerulonephritis; vasculitis; nephritis;
KM renal neoplasm; multiple myeloma; lymphoma; light chain neuropathy;
KM immune response; immunosuppression; graft rejection; joint pain;
KM graft versus host disease; inflammation; swelling; anaemia; septic shock;
KM insulin dependent diabetes mellitus; Crohn's disease; hypertension;
KM renal artery stenosis; occlusion; cholesterol; renal emboli.

Mus musculus

PN W0200040716-A2

PD 13-JUL-2000

07-JAN-2000; 2000WO-US00396.

PR 07-JAN-1999; 99US-0226533.

PA (ZYMO) ZYMOGENETICS INC.

PI Gross JA, Xu W, Madden K, Yee DP;

DR WPI; 2000-452538/39.

Tobias - 4-6A
XX
XX

Disclosures: Page 163, 175pp; English.

PS Disclosure; Page 163; 175pp; English

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:48:20 ; Search time 14.98 seconds
(without alignments)
132.074 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 498

Sequence: 1 CSQNEYPDSLHACIPQCLR.....SEYFDSLHACIPATCPYC 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issuel_Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pcp:*

2: /cgn2_6/ptodata/2/1aa/5B.COMB.pcp:*

3: /cgn2_6/ptodata/2/1aa/6A.COMB.pcp:*

4: /cgn2_6/ptodata/2/1aa/6B.COMB.pcp:*

5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pcp:*

6: /cgn2_6/ptodata/2/1aa/Backfile1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	18.7	166	2	US-08-810-572A-6
2	93	18.7	166	2	US-09-290-333-6
3	93	18.7	293	2	US-08-810-572A-2
4	93	18.7	293	4	US-09-290-333-2
5	82	16.5	508	4	US-09-019-095A-8
6	82	16.5	521	4	US-09-019-095A-22
7	82	16.5	526	4	US-09-019-095A-2
8	80	16.1	3075	2	US-08-460-309-5
9	80	16.1	3075	2	US-08-125-077-5
10	80	16.1	5405	4	US-08-718-388-9
11	79	15.9	545	4	US-09-019-095A-38
12	78	15.7	109	2	US-08-527-044-2
13	78	15.7	139	3	US-09-013-780-2
14	77.5	15.6	109	3	US-08-965-903B-20
15	76.5	15.4	2414	1	US-08-227-536-2
16	76.5	15.4	2414	1	US-08-965-903B-20
17	76	15.3	2441	1	US-08-194-468-2
18	76	15.3	2441	3	US-08-961-739-2
19	75.5	15.2	156	3	US-08-600-982-30
20	75.5	15.2	156	5	PCT-US94-10261A-30
21	75.5	15.2	1713	3	US-08-600-982-24
22	75.5	15.2	1713	3	PCT-US94-10261A-24
23	75	15.1	219	1	US-08-152-019A-31
24	75	15.1	219	2	US-08-460-309-18
25	75	15.1	219	2	US-08-125-077-18
26	75	15.1	430	3	US-08-997-897-2
27	75	15.1	430	4	US-09-156-836B-2

28	73	14.7	689	4	US-09-177-249-2	Sequence 2, Appl
29	73	14.7	689	4	US-09-061-769A-2	Sequence 2, Appl
30	72.5	14.6	969	2	US-08-284-941-2	Sequence 2, Appl
31	72.5	14.6	969	2	US-08-447-642-2	Sequence 2, Appl
32	72.5	14.6	969	2	US-09-236-503-2	Sequence 2, Appl
33	72.5	14.6	969	5	PCT-US93-02147A-2	Sequence 2, Appl
34	72	14.5	275	1	US-08-312-870-7	Sequence 7, Appl
35	72	14.5	366	4	US-08-857-076-103	Sequence 103, App
36	72	14.5	446	1	US-08-307-444A-5	Sequence 5, Appl
37	72	14.5	446	1	US-08-587-389-5	Sequence 3, Appl
38	72	14.5	456	1	US-08-307-444A-3	Sequence 3, Appl
39	72	14.5	456	1	US-08-307-444A-4	Sequence 4, Appl
40	72	14.5	456	1	US-08-587-389-3	Sequence 3, Appl
41	72	14.5	456	1	US-08-587-389-4	Sequence 4, Appl
42	72	14.5	475	1	US-08-307-444A-1	Sequence 1, Appl
43	72	14.5	475	1	US-08-307-444A-2	Sequence 2, Appl
44	72	14.5	475	1	US-08-587-389-1	Sequence 1, Appl
45	72	14.5	475	1	US-08-587-389-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-810-572A-6
; Sequence 6, Application US/08810572A
; Patent No. 5969102
; GENERAL INFORMATION:
; APPLICANT: Bram, Richard J.
; TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS
; TITLE OF INVENTION: CAM1, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,572A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ. ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-810-572A-6

Query Match 18.7%; Score 93; DB 2; Length 166;

SEQ ID NO 2
LENGTH: 526
TYPE: PRT
ORGANISM: Murine
US-09-019-095A-2

Query Match 16.5%; Score 82; DB 4; Length 526;
Best Local Similarity 35.8%; Pred. No. 1.1;
Matches 19; Conservative 8; Mismatches 18; Indels 8; Gaps 2;

QY 19 LRCSSNTPPLTCQRCCEYFDSLHACPLRCSPTQCYCCFHSYFDSLHA 71
Db 66 LQCLHTRPPL-----ADYMLSOEHSQTC--CSPEGCKLCAMEALVTSLLHS 110

RESULT 8
US-08-460-309-5
Sequence 5, Application US/08460309
Patent No. 5837496
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460.309
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/125.077
FILING DATE: 22-SEP-1993
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472.319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919.951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3075 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-460-309-5

Query Match 16.1%; Score 80; DB 2; Length 3075;
Best Local Similarity 23.4%; Pred. No. 9;
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;
QY 1 CSQNEYFDSL--HACIPQ-----LRCSSNTPPLTCQRCCE-- 36

Db 831 CADGYGNPTVPGESCVCDSGNVDPSEAGHCDVTGECIKLGNTDGAHCR-CADGF 889
QY 37 YFDSL-----LHACPC-----LRCSP--PTQCYC--CFHSEYFDSLHACP 73
Db 890 YGDVATKNCRCRCHVKGSHAVCHLETGLCDCKPNTVG00CDQCLHGYGLDGSgCCR 949
QY 74 PATC 77
Db 950 PCNC 953

RESULT 9
US-08-125-077-5
Sequence 5, Application US/08125077
Patent No. 5872231
Patent No. 5872231 5840863
GENERAL INFORMATION:
APPLICANT: Engvall, Eva
APPLICANT: Leivo, Ilmo
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin
TITLE OF INVENTION: Fragments and Uses Thereof
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/125.077
FILING DATE: 22-SEP-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US 94/10730
FILING DATE: 21-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/472.319
FILING DATE: 30-JAN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/919.951
FILING DATE: 27-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LA 9721
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3075 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-125-077-5

Query Match 16.1%; Score 80; DB 2; Length 3075;
Best Local Similarity 23.4%; Pred. No. 9;
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;

QY 1 CSQNEYFDSL--HACIPQ-----LRCSSNTPPLTCQRCCE-- 36
Db 831 CADGYGNPTVPGESCVCDSGNVDPSEAGHCDVTGECIKLGNTDGAHCR-CADGF 889
QY 37 YFDSL-----LHACPC-----LRCSP--PTQCYC--CFHSEYFDSLHACP 73

Db 89 YGDAVTAKMCRACECHVKGSHSAVCHLETGLCDCKBNVTGQDCDOCLHGYYGLDSHGCR 949
QY 74 PATC 77
Db 950 PCNC 953

RESULT 10
US-08-718-388-9
; Sequence 9, Application US/08718388
; Patent No. 6271362
; GENERAL INFORMATION:
; APPLICANT: MORIKAWA, MINORU
; APPLICANT: HARADA, NAOKI
; TITLE OF INVENTION: GENE ENCODING IGG FC REGION-BINDING
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,388
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: MURPHY JR, GERALD M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 0230-111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5405 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-718-388-9

Query Match 16.1%; Score 80; DB 4; Length 5405;
Best Local Similarity 25.4%; Pred. No. 15;
Matches 29; Conservative 12; Mismatches 25; Indels 48; Gaps 10;

QY 1 CGSNEYFDSLHKCI-PCQLRCSNTPPLTCQRYCCCFPSLHACPC-----LRCS 51
Db 2733 CPONSIVE-----CADTCSLGSALSAPLQCPDGCAC-----GQCDSGFLYNGQACV 2781

QY 52 PPTCQYRC-CFHS-EYFD---SLT-----HACPPA-TCOP 79
Db 2782 P--IQCCGCHNCAYYEPQYVLIDMCRQOCCTHAGKYVVCQHSCKPQVQCP 2833

RESULT 11
US-09-019-095A-38
; Sequence 38, Application US/09019095A
; Patent No. 6287858
; GENERAL INFORMATION:
; APPLICANT: D'Andrea, Alan D.
; APPLICANT: Zhu, Yuan
; TITLE OF INVENTION: Deubiquitinating Enzymes That Regulate
; TITLE OF INVENTION: Cell Growth
; FILE REFERENCE: DEPT-43502A2
; CURRENT APPLICATION NUMBER: US/09/019,095A

; CURRENT FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: PCT/US96/12834
; PRIOR FILING DATE: 1996-08-07
; PRIOR APPLICATION NUMBER: US 60/002,066
; PRIOR FILING DATE: 1995-08-09
; PRIOR APPLICATION NUMBER: US 60/019,787
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 545
; TYPE: PRT
; ORGANISM: murine
US-09-019-095A-38

Query Match 15.9%; Score 79; DB 4; Length 545;
Best Local Similarity 34.0%; Pred. No. 2.1;
Matches 18; Conservative 9; Mismatches 18; Indels 8; Gaps 2;

QY 19 LRCSNTPPLTCQRYCCCFPSLHACPCRLGSPPTQYCCFHSYFDSLHA 71
Db 66 LQCLTHPLPPL-----ADYMLSOBYSTQC--CSPEGCKMCAEHAHTQSLHS 110

RESULT 12
US-08-527-044-2
; Sequence 2, Application US/08527044
; Patent No. 5885568
; GENERAL INFORMATION:
; APPLICANT: BUNSTEAD, Janene Marilyn J.M.
; APPLICANT: TOMLEY, Fiona Margaret F.
; APPLICANT: DUNN, Patrick Paul James P.
; APPLICANT: VERMEULEN, Arnoldus Nicolaas A.N.
; TITLE OF INVENTION: Coccidiosis poultry vaccine
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo No. 5885568e1 N.V
; STREET: 1300 PICCARD DRIVE #206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850-4373
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/527,044
; FILING DATE: 12-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 94202667.6
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GORMLEY, MARY E.
; REGISTRATION NUMBER: 34,409
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-527-044-2

Query Match 15.7%; Score 78; DB 2; Length 109;
Best Local Similarity 26.7%; Pred. No. 0.58;
Matches 20; Conservative 7; Mismatches 24; Indels 24; Gaps 5;

QY 10 LHAICPQLRCSNTPPLTCQRYCCCFPSLHACPCRLGSPPTC---QYCCFHSYF 66
Db 6 LTNAC--CCLRYTNS-----CCSKYCS-----KCCCKKCCSKCCCKCTYCC----- 45

QY	67	SLHACPRATCQPYC	81
			:
Db	46	-STFCCSKCCCSKFC	59

RESULT 13
US-09-013-780-2
; Sequence 2, Application US/09013780
Reference: 508133

GENERAL INFORMATION: BUNSTREAD, Janene Matilyn J.M.
APPLICANT: TOMLEY, Fiona Margaret F.
APPLICANT: DUNN, Patrick Paul James P.
APPLICANT: VERMEULEN, Arnoldus Nicolaas A.N.
TITLE OF INVENTION: Coccidiosis poultry vaccine
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/013,780
FILING DATE:

Query Match	15.7%	Score	78	DB	3	Length	109
Best Local Similarity	26.7%	Pred. NC	0.58				
Matches	20	Conservative	7	Mismatches	24	Indels	24
						Gaps	5

QY 10 LTHACIPCOLRCSNNPPLTCQRYCCETFDLSLHACPCLRKSPPTC---QYCCFHSFYD 66
| : | | : : | : | : | : |
Db 6 LTNMC--CCRLRTNS---CCSKRYCCS-----KCCSCCKCCSKCCSTYCC----- 45

```

QY      67 SLHACPRATCQPYC 81
          |      | :|
Db      46 -STFCCSKCCCKKFC 59

```

RESULT 14
US-08-965-903B-20
Sequence 20, Application US/08965903B
Patent No. 6060275
GENERAL INFORMATION:
APPLICANT: Hachsen, Nit
TITLE OF INVENTION: SPROUTY PROTEIN AND CODING
TITLE OF INVENTION: SEQUENCE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Delinger & Associates
STREET: 350 Cambridge Ave., Suite 250

```

1 CITY: Palo Alto
2 STATE: CA
3 COUNTRY: USA
4 ZIP: 94306
5
6 COMPUTER READABLE FORM:
7 MEDIUM TYPE: Diskette
8 COMPUTER: IBM Compatible
9 OPERATING SYSTEM: DOS
10 SOFTWARE: FASTESTO for Windows Version 2.0
11
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/965,903B
14 FILING DATE: 07-NOV-1997
15
16 CLASSIFICATION: 435
17
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 60/030232
20 FILING DATE: 07-NOV-1996
21
22 ATTORNEY/AGENT INFORMATION:
23 NAME: Petithory, Joanne R
24 REGISTRATION NUMBER: 42,995
25 REFERENCE/DOCKET NUMBER: 8600-0177.30
26
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: 650-324-0880
29 TELEFAX: 650-324-0960
30
31 TELEX:
32
33 INFORMATION FOR SEQ ID NO: 20:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 139 amino acids
36 TYPE: amino acid
37 STRANDEDNESS: single
38 TOPOLOGY: linear
39 MOLECULE TYPE: protein
40 FRAGMENT TYPE: internal
41
42 ORIGINAL SOURCE:
43 INDIVIDUAL ISOLATE: h-spryl
44
45 US-08-965-903B-20

```

Query Match	15.68;	Pred	77.5;	DB	3;	Length	139;
Best Local Similarity	30.38;	Pred. No.	0.81;				
Matches	27;	Conservative	8;	Mismatches	29;	Indels	25;
						Gaps	6

```

0y 1 CSQNEYFDLSLHACIPQLKRSSNTPPLTQRCCEYFDLSLHACIPQLKRCSPPT----- 54
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 52 CSNDEGDS--YSDNRCS--CSQSH---CCSRYLICMGAMSLF--LPCLLCYPPAKGCLKL 102

```

```

QY      55  CQYC-----CFHSEYFDLSLHACP  73
          | : |           | : |
Db      103  CRRCYDWIHRPGCRCKNSNTVYCKLESCP  131

```

RESULT 15
 US-08-227-536-2
 Sequence 2, Application US/08227536
 Patent No. 5658784
 GENERAL INFORMATION:
 APPLICANT: Eckner, Richard
 APPLICANT: Ewen, Mark
 APPLICANT: Livingston, David
 TITLE OF INVENTION: NOCLEIC ACID, ENCODING TRANSCRIPTION
 TITLE OF INVENTION: FACTOR P300 AND USES OF P300
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
 STREET: Ten Post Office Square
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25

